



The Alberta Journal of Educational Research

Volume XIX, No. 1

March, 1973

CONTENTS

Reading Comprehension Tests: Multiple Guess?	1
<i>Robert F. Fischer</i>	
An Abbreviated Version of the 1960 Stanford-Binet, Form L-M	7
<i>Frances Layton</i>	
Trends in the Patterns of Staff Utilization in the Elementary Schools of Alberta: Implications for Teacher Education	12
<i>R. D. Armstrong and N. M. Purvis</i>	
Political Climates, Political Knowledge and Political Preferences of University Students	22
<i>Farhat Ghaem-Maghami</i>	
Selecting Instructional Materials to Meet Students' Reading Performance Level	30
<i>Clifford D. Pennock</i>	
The Effects of Human Relations Training on Interpersonal Skills	37
<i>John P. Miller</i>	
Teacher Attitude as a Function of Locus of Control	48
<i>Henry L. Janzen, Don Beeken and John Hritzuk</i>	
Canadian School Life Tables	55
<i>Jim Stafford</i>	
Teacher Education: Environments and Outcome	66
<i>John McLeish</i>	
Reference Books Received	82
<i>The Editor</i>	

FACULTY OF EDUCATION
The University of Alberta

ROBERT F. FISCHER

Student Counselling Services
The University of Alberta

Reading Comprehension Tests: Multiple Guess?

This study was designed to determine whether certain commercial 10-item multiple-choice reading comprehension tests were vulnerable to guessing. Forty university students were given tests at an "Easy" level, and 40 at a "Difficult" level, without having read the material on which the tests were based. Mean scores were 5.78 for two Easy tests and 5.23 for two Difficult ones, compared with predicted chance scores of 2.5 and 2.0 respectively. Some individual items were guessed correctly by all subjects, and others were guessed incorrectly by almost all. It was concluded that more valid comprehension tests could be developed if items were tested for guessability.

Students commonly refer to multiple-choice tests as "multiple-guess," and there is some evidence that in reference to so-called "reading comprehension" tests, the put-down may be justified. Carver (1972) has criticized speed-reading courses as exercises in skimming rather than reading, and one of his main criticisms is that the tests used to prove that students in these courses maintain high levels of comprehension are invalid. He found that subjects were able to score an average of 57% on a "comprehension" test without having read the material on which it was based. A search of the psychological literature of the years 1967 to 1972 yielded only one reference to the general subject (Samuels, 1968) in addition to Carver's article, and in this case the factor under study was the effect of the wording of items on subjects' guesses, rather than the effect of guessing on test scores.

The present study was designed to find out whether tests used in one of the popular do-it-yourself reading-improvement kits available in many academic counselling settings would be vulnerable to educated guessing.

Method

Subjects of the study were 80 students (50 male, 30 female) or prospective students who were seen individually by the experimenter in the Student Counselling Services at The University of Alberta. Over a period of several weeks each of the experimenter's counselees (most of them being seen for the first time) who fit certain criteria were asked if they would be willing to spend 10 minutes at the end of the hour as subjects in a research project. Criteria for selection were that the subject be either a student enrolled in the university or a Grade 12 graduate fully qualified for admission, and that he or she be relatively free of serious emotional disturbance as judged by the experimenter. Subjects were thereby limited to clients who presented vocational and/or academic problems.

Each subject performed three tasks: (a) read a selection, (b) wrote a comprehension test on the selection, and (c) wrote a comprehension test on a selection that had not been read.

Reading selections were taken from the Science Research Associates Reading Laboratory Kit, 1959 Collège Prep Edition. Materials in this kit are color-coded at seven levels of difficulty. For this experiment two selections were chosen at each of two levels: the "Easy" selections were "Olive" numbers 5 and 8 (the second easiest level), and the "Difficult" selections "Purple" numbers 7 and 10 (the most difficult level). The Easy selections will be referred to as A and B, and the Difficult selections as C and D.

Selections A and B contained approximately 250 words each, and C and D approximately 400. Subjects were allowed 90 seconds to read A or B and three minutes to read C or D. It was assumed that under these limits each subject should be able to read the selection at least once. They were informed of the time limits and were instructed to continue studying the material if they finished before time was up. Reading selections and tests were typed on separate sheets so that there could be no re-reading of the material after the test was presented.

At each level of difficulty, subjects were assigned to four treatment groups, consecutively in the order in which they appeared for counselling, according to the following scheme:

1. Read A, test on A, test on B
2. Test on B, read A, test on A
3. Read B, test on B, test on A
4. Test on A, read B, test on B

followed by a similar pattern for C and D.

This pattern was designed to control for position effects and differences in difficulty between the two tests at each level. Half of the subjects in each group would read the selection first, and half would write the "cold" test first. Furthermore, half of the subjects in the "Easy" group would write test A after having read A, and would write test B "cold"; half would write test B after having read B, and would write test A "cold," and so on. There were 10 subjects in each of the eight treatment groups.

Each test was made up of 10 multiple-choice items. Items on tests A and B offered four response choices, and items on tests C and D offered five choices.

Data were analyzed separately for each level of difficulty. Mean test scores for each of the eight treatment groups were calculated. To obtain means for performance after reading ("Read" scores) and performance without reading ("Non-read" scores), scores for the appropriate treatments on tests A and B and on tests C and D were combined.

In addition, the frequencies of all obtained scores were compared with the frequencies expected by chance on the 4-choice and the 5-choice tests.

Individual items were also analyzed. To find out whether there were differences among the items in terms of difficulty or "guessability," the total numbers of correct responses for each item under "Read" and "Non-read" conditions were determined.

Results

Mean test scores for each of the eight treatment groups and for combinations of the groups are presented in Table 1. By inspection it appears that there were no important differences between tests administered first and second in sequence, and therefore no significant position effects. At the Easy level tests A and B were of approximately equal difficulty for the Readers, but test B seems to have been easier for the Guessers, who made a mean score of 6.85 on test B and 4.70 on test A. At the Difficult level there was no difference in guessability between tests C and D, but C seems to have been a little easier as a comprehension test for the Readers.

TABLE 1
MEAN SCORES ON COMPREHENSION TESTS FOR FOUR READING SELECTIONS

Selection	Read			Non-Read		
	First	Second	Total	First	Second	Total
A	7.60	8.10	7.85	4.80	4.60	4.70
B	8.20	8.10	8.15	7.10	6.60	6.85
A + B	7.90	8.10	8.00	5.95	5.60	5.78
C	7.10	7.50	7.30	5.00	5.40	5.20
D	6.40	6.90	6.65	5.30	5.20	5.25
C + D	6.75	7.20	6.98	5.15	5.30	5.23

When the total scores for the 40 subjects at each level were considered, the most significant finding was that the Guessers performed quite well. At the Easy level their mean score was 5.78, and at the Difficult level 5.23. Since the number of response choices on each item was four at the Easy level and five at the Difficult level, the mean chance scores on the 10-item tests would be 2.5 and 2.0, respectively. Mean scores of the Readers were 8.00 at the Easy level and 6.98 on the Difficult tests.

Obtained frequencies of scores on the "Read" and "Non-read" tests are presented in Table 2.

TABLE 2
EXPECTED AND OBTAINED FREQUENCIES OF SCORES ON READ AND
NON-READ COMPREHENSION TESTS

Selections A & B				Selections C & D		
Score	Expected Frequencies ^a	Obtained Frequencies		Expected Frequencies ^a	Obtained Frequencies	
		Read	Non-Read		Read	Non-Read
0	2.25	0	0	4.29	0	0
1	7.51	0	0	10.74	0	0
2	11.26	0	1	12.08	0	2
3	10.01	1	2	8.06	0	2
4	5.84	0	7	3.52	2	7
5	2.34	2	7	1.06	5	9
6	0.65	4	9	0.22	5	14
7	0.12	8	7	0.03	12	6
8	0.02	6	6	b	13	0
9	b	11	1	b	2	0
10	b	8	0	b	1	0

^a Frequencies expected by chance in 40 cases

^b Negligible

Expected frequencies were determined in a manner similar to the following example: on a 10-item test, three correct answers can be obtained in 120 ($10 \times 9 \times 8$ divided by $1 \times 2 \times 3$) different ways; if there are four response choices the remaining seven items can be answered incorrectly in 3^7 different ways; and the total number of possible combinations of answers is 4^{10} . Therefore the probability of obtaining three correct answers by chance is 120×3^7 over 4^{10} , or .2503, and this figure multiplied by 40 gives the expected frequency of 10.01 for 40 subjects.

It is obvious by inspection that chi-square tests applied to the differences between expected and obtained frequencies for both the Read and the Non-read tests would yield values of astronomical magnitude. In only one case in 80 did a Guesser score below the chance level—a score of 2 on test A. Two other Guessers made precisely chance scores of 2, one on each of tests C and D. Whereas the expected frequency of scores of 5 or higher on the two Difficult tests was 1.31 in 40 cases, the obtained frequency of such scores was 29. At the Easy level 30 of the 40 subjects scored 5 or higher, in contrast to an expected frequency of 3.13. It seems safe to conclude that as a group the Guessers scored significantly higher than would be expected by chance.

Obtained frequencies of correct answers on each item of the four tests are given in Table 3. The maximum possible score for each item of course would be 20, since each test was written by 20 subjects. Differences in guessability among the items were quite striking, as shown by the extreme

range of the Non-read scores—from zero to 20. At one extreme were four items that were answered correctly by all of the Guessers—and in each case answered incorrectly by one or more of the Readers. At the opposite extreme were a number of items that seemed to be very difficult to guess correctly but comparatively easy to answer after reading—for example items 2 and 4 on test A, and item 6 on test B.

TABLE 3
FREQUENCIES OF CORRECT ANSWERS ON EACH ITEM OF FOUR
COMPREHENSION TESTS

Item	Test A		Test B		Test C		Test D	
	Read	Non-Read	Read	Non-Read	Read	Non-Read	Read	Non-Read
1	16	20	16	15	18	16	15	16
2	12	0	14	9	18	14	7	9
3	19	10	15	18	16	7	19	20
4	18	3	10	14	13	10	19	11
5	15	12	20	16	11	10	10	6
6	14	5	17	3	18	5	7	7
7	17	10	13	12	14	7	18	16
8	17	8	20	19	17	20	16	5
9	18	20	20	16	17	11	15	13
10	11	6	18	15	4	4	7	2

Discussion

In this study students were able to score from 47% to 68% on commercial “reading comprehension” tests without reading the material on which the tests were based. It is obvious that they were doing more than mere guessing—drawing on past learning, picking up cues from one item to answer another, critically weighing the wording of items, etc.

The most significant finding of the study would seem to be that tests and individual items varied widely in their vulnerability to “guessing.” In the tests used in this study there were four items that were answered correctly by all of the Guessers. Two of these items called for the definition of single words, and the other two were expressed in such terms that any moderately sophisticated university student might be expected to find only one answer reasonable. (It may be noted with some gratification that one of the words that was recognized by all of the student subjects was “democracy.” The other, alas, was “archipelago.”) The fact that some items proved almost invulnerable to guessing suggests that more valid comprehension tests could be developed if items were tested for guessability before being included. Samuels’s (1968) finding that guessers were influenced in their choices by certain word associations seems pertinent to this problem.

Another way of expressing the findings of this study would be to state that the SRA reading kit tests are not very reliable, because they are

highly susceptible to guessing; and of course their low reliability is also due in part to their brevity. The phenomena of brevity, guessability, and low reliability go hand-in-hand-in-hand. As Cronbach (1960) points out in a discussion of reliability, "If a test has only five multiple-choice items, a few people might get all the items correct just by guessing. In a fifty-item test practically no one could do well by guessing. Variations due to guessing tend to cancel out [p. 130]." There is no technical manual accompanying the SRA kit, and hence no information on reliability, and it seems unlikely that the tests (140 in all) have been subjected to reliability studies.

The manual for the reading accelerator that is used with the SRA materials (Simpson, 1958) suggests that the student has increased his paced speed by 15% if "the average of the last two comprehension scores is 70% or better [p. 16]." The assumption seems to be that a score of 70% indicates adequate comprehension. In general practice there seems to be an implication that the test score gives a measure of the quantity of comprehension—that a score of 50%, for example, means that the reader has remembered 50% of the material. It appears more likely that a score of 50% means nothing at all.

References

- Carver, R. P. Speed readers don't read; they skim. *Psychology Today*, August, 1972, 6, 22.
- Cronbach, L. J. *Essentials of psychological testing*. (2nd ed.) New York: Harper & Row, 1960.
- Samuels, S. J. Effect of word associations on reading speed, recall, and guessing behavior on tests. *Journal of Educational Psychology*, 1968, 59 (1, Pt. 1), 12-15.
- Science Research Associates, Inc. Reading Laboratory IVa, College Prep Edition, 1959.
- Simpson, E. A. *How to use the SRA Reading Accelerator*. Chicago: Science Research Associates, Inc., 1958.

FRANCES LAYTON

University of Oregon

An Abbreviated Version of the 1960 Stanford-Binet, Form L-M

The purpose of this study was to test a short version of the Stanford-Binet, Form L-M using a group covering a wide age and ability level. The prime reason was to attempt to reduce the time factor involved in administration of some of the S-B tests, without sacrificing the reported accuracy. From the results it appears that this short method of S-B administration will reduce the time needed to administer the S-B without adversely affecting the accuracy of the I.Q. score.

Little research is available concerning the Short Form method of administering the Stanford-Binet Form L-M, published in 1960. Investigations with the short versions have been carried out with specialized groups, usually with mentally retarded and occasionally with disadvantaged Negro groups. Employing either the short form according to the manual (Terman and Merrill, 1960) or the procedure designed by Wright (1942), studies have shown correlations from .93 to .99 with the unabbreviated Stanford-Binet.

Although the S-B is a proven individual test of mental ability for school age children, it is used with reluctance in many cases because of the time factor involved. It often occurs that a child will continue to score on one sub-test at various age levels before a ceiling age can be determined; the reverse, when attempting to determine the basal age, is also true. There appears to be some merit, therefore, in investigating the possibility of employing a shorter version of the S-B for routine individual intelligence testing, especially in the schools. The question arises concerning the advisability of investigating this possibility using mentally retarded or specialized groups. A strong thesis could be formed in favor of extending and lengthening the method used in testing these groups and not in shortening it.

In the S-B manual (Terman and Merrill, 1960) the method employed to shorten administration is to use four selected tests at each age level

and to weight these tests more heavily than with normal scoring (p. 61). The suggested time saving is 20 per cent.

Wright's method (1942) of S-B abbreviation is slightly different. She advocates administering all six original items at the levels at which the subject passes, or fails, the four items in the recommended short version. To attain a basal age or ceiling age, all six items at the next, lower or higher, levels, are administered if necessary. Robinson and Robinson report that this method saves relatively little time and the longer form is preferable (1965, p. 413).

The proposition is made here that the recommended short version of the S-B, and Wright's version, both have the fault that they use only a sample of the test items and difficulties can arise from this. In studies reported by Watson (1963) both versions result in lower mean I.Q.'s and occasional wide discrepancies from the fully administered S-B. Wright's version also has a disadvantage in that it appears to save little time.

It is suggested here that the criteria for determining basal and ceiling ages be as follows: basal age is that age level at which all sub-tests are passed but one; ceiling age is that age level at which all sub-tests are failed but one. This method would make use of *all* items at age levels and not a sampling of those items. The rationale for this suggestion is that it is not unusual that a testee passes or fails all but one sub-test at an age level, thus extending the time required for the test.

Method

Five hundred S-B protocols, covering an age range from three years four months to seventy years six months and an I.Q. range from 160 to 45 were investigated. These protocols were from a wide variety of sources and not limited to referral or clinical cases, although some of these were included.

The protocols were re-scored employing the following methods. With the individuals whose chronological age was six years or over, the vocabulary section dictated the age level at which the scoring would begin. With individuals whose ages were under six years the starting point was one level below that of their chronological age. Basal age was determined as the age level at which all sub-tests were passed but one. Ceiling age was determined as that level at which all sub-tests were failed but one.

Tests were re-scored in this manner and any change in mental age and I.Q. was recorded. The number of levels omitted because of the new scoring procedure was recorded for each case.

A test of significance of difference between the mean I.Q. of the fully administered S-B and the mean I.Q. of the short version was used. Pearson product-moment correlation was calculated for the relationship between the fully administered and the employed short method. This analysis was carried out with the total number of tests and again with those tests which showed a difference in I.Q. points because of the shortened version of the S-B.

Results

Table 1 indicates that when the 500 tests were re-scored using the method described, 611 age levels were omitted or saved; 203 tests showed no I.Q. change even with 260 age levels omitted; 126 tests showed an I.Q. difference with 351 age levels omitted; and the short method was not useful with 171 of these tests, the time factor being the same by both methods.

TABLE 1
RESULTS OF RE-SCORING THE 500 TESTS

Description	No. Tests	Levels Omitted
No change in I.Q.	203	260
Differences in I.Q.	126	351
Short version not useful	171	0
Total	500	611

Table 2 indicates the differences between the fully administered tests and the shortened method of scoring.

TABLE 2
DIFFERENCES BETWEEN THE STANDARD AND
SHORTENED SCORING METHODS

Tests	N	Levels Omitted	Mean I.Q.	S.D. I.Q.	C.R.
Standard scoring	500	—	114.192	20.05	.012
Short version	500	611	114.176	20.71	

With the 126 tests which showed a difference in I.Q. score, 351 levels were omitted and there was a 1.27 difference between the means of the two groups. Table 2 shows the comparison between the two groups.

TABLE 3
COMPARISON OF STANDARD SCORING AND SHORTENED METHODS

Tests	N	Levels Omitted	Mean I.Q.	S.D. I.Q.	C.R.
Standard scoring	126	—	113.71	23.28	.603
Short version	126	351	112.44	23.28	

The overall difference in I.Q. points between the 500 fully administered and short version tests amounted to .996 I.Q. points per test with 1.22

age levels saved per test. With the 126 tests showing a difference of I.Q. points, the change was 3.96 I.Q. points per test and 2.84 levels omitted per test.

Over the total number of tests (500), the correlation was .9592 between the fully administered S-B and the short version. A correlation of .9991 was obtained between the 126 tests which had shown a difference in I.Q. points.

Discussion and Conclusions

This short version has the advantage that it does not sample subtests from age levels as recommended by the S-B manual (Terman and Merrill, 1960), and recommended by Wright (1942). According to Robinson and Robinson (1965) the method proposed by Wright saves little time in actual test administration.

The wide variety of age and ability levels was selected because it was felt that an investigation of this type must first be carried out with a group other than a mentally retarded or exceptional group. The reasons for this decision were presented earlier in this paper.

While this short version compares favorably with the full scale administration of the S-B and eliminates a considerable amount of testing time, there are several cautions needed. Any method of data analysis tends to conceal anomalies within the data. Five of the 500 cases showed an I.Q. difference varying from 9 to 12 points. One case differed by 18 I.Q. points. These six cases involved children under the chronological age of 11-3 years and with a superior or above average mental age level. This might be explained as a function of some age levels but this investigation must be replicated in an effort to explain these changes. Notwithstanding, this short version appears to have some merit without the disadvantages of the other two procedures discussed. It must be considered too, that most previous studies using abbreviated versions of the S-B have been done with mentally retarded groups. These groups do not appear to suffer I.Q. changes when a short method is employed. In this study, if cases with I.Q.'s less than one standard deviation below the mean are isolated, there are no I.Q. changes more than three points; two cases lost three I.Q. points, one case gained three points and one gained two points. Fourteen cases remained unchanged.

Before any abbreviated version of the S-B can be safely employed, it must first be fully investigated using age and ability levels encompassing a wide range. The present study suggests a method of abbreviating the time taken in testing using an *ex post facto* analysis. The usefulness of the abbreviation suggested still requires experimental validation.

References

- Kennedy, W. A., Van de Reit, V., & White, J. C. Use of the Terman-Merrill abbreviated scale of the 1960 Stanford-Binet Form L-M on Negro elementary school children in the southeastern United States. *Journal of Consulting Psychology*, 1963, 27, 456-457.
- Robinson, H. B., & Robinson, N. M. *The mentally retarded child*. New York: McGraw, 1965.
- Silverstein, A. B. The evaluation of two short forms of the Stanford-Binet for use with mentally retarded children. *American Journal of Mental Deficiency*, 1963, 67, 922-923.
- Terman, L. M., & Merrill, M. A. *Measuring intelligence*. Boston: Houghton Mifflin, 1937.
- Terman, L. M., & Merrill, M. A. *Stanford-Binet intelligence scale: Manual for the third revision form L-M*. Boston: Houghton Mifflin, 1960.
- Watson, R. I. *The clinical method in psychology*. New York: Harper, 1951.
- Wright, C. A. A modified procedure for the abbreviated Stanford-Binet in determining the intelligence of mental defectives. *American Journal of Mental Deficiency*, 1942, 47, 178-184.

R. D. ARMSTRONG

University of Victoria

and

N. M. PURVIS

The University of Alberta

Trends in the Patterns of Staff Utilization in the Elementary Schools of Alberta: Implications for Teacher Education

While teacher education institutions must lead as well as follow, accurate information as to the status of the various instructional and organizational practices in the schools is certainly one prerequisite for the planning of both preservice and in-service curricula. In reporting on a survey of team teaching practices in western Canada, MacKay and Ward (1967) conclude, in part: "... those who are concerned with educational decision making in Canada must take account of the effect such developments have upon staffing needs, building design, in-service education, and teacher education programs (p. 12)." The purpose of the study reported here was to examine the patterns of staff utilization in the elementary schools of Alberta in 1968 and in 1971 and to draw implications for teacher education.

During the last ten years, team teaching and other patterns of staff utilization resulting in a move away from the self-contained classroom have received increasing attention both in the literature and in practice. In western Canada, for example, MacKay and Ward (1967) reported that, "the practice of team teaching is fairly extensive," while one year later Ratsoy (1968) concluded: "Team teaching is growing rapidly in the West." (p. 49)

Definitions of the four patterns of staff utilization examined in the present study appear below. The definitions appear defensible but, obviously, not all team teaching, as an example, is the same. Anderson (1965) has suggested that "dozens of different patterns of school and staff organization," (p. 22) may be found under the general heading of co-operative teaching. Shaplin, in Shaplin and Olds (1964), in writing about

team teaching projects, remarks: “. . . one is impressed by their great diversity in methods of organization and aims.” (p. 5)

The location of an elementary school, whether in a rural or an urban setting, would seem to justify certain expectations in terms of staff utilization. In a survey of practices in music education in Alberta, Coultas (1965) noted important differences in urban and rural systems.

In cities and towns approximately one-tenth of classroom teachers were responsible for teaching music as compared with approximately two-thirds of teachers in divisions and counties. (p. 96)

Blench's survey of the teaching of science in Alberta, reported in 1967 that many urban elementary schools were using a type of departmentalization or semi-departmentalization to teach science while in rural schools science was more likely to be taught by the classroom teacher (p. 133). Also of interest to the present study is Blench's report of the contrast between primary and intermediate grades in terms of teacher specialization (p. 133).

Many studies, including the present one, report dramatic changes in patterns of staff utilization as well as marked increases in the relative amount of team teaching. The results of these studies should probably be considered in the context of recommended procedures for introducing team teaching. Polos (1965), as a typical example, strongly supports the need for in-service training to precede and accompany the move toward team teaching (p. 107).

Procedures

The four patterns of staff utilization investigated were defined as follows:

1. *Team or Cooperative Teaching*: One or more teams in which two or more teachers per team jointly plan, present, and evaluate in-school experiences for pupils from grade one to grade six inclusive.
2. *Departmentalized Teaching*: No team or cooperative teaching but with 50 per cent or more of the teachers' time, on the average, spent in teaching classes other than their own.
3. *Semi-Departmentalized Teaching*: No team or cooperative teaching, some exchange of subjects but with 50 percent or more of the teacher's time, on the average, spent in teaching her own class.
4. *Self-Contained Classrooms*: No team or cooperative teaching, with teachers teaching all subjects in their own classrooms.

In 1968 all superintendents in the province of Alberta and eighty-three elementary school principals and their staffs were asked to fill out a questionnaire to determine the patterns of staff utilization being used in the elementary schools of the province and the preferences of the various personnel for these patterns. The principals and school staffs were selected as a random sample, stratified to represent the rural-urban distribution of schools, schools of different sizes, and the four different staff utilization patterns as indicated by the superintendent. All data to follow were obtained from the superintendents except for the preferences of

the principals and teachers. All superintendents returned the questionnaires and approximately 86 per cent of the school staffs.

In 1971 all superintendents were again asked to report on the patterns of staff utilization in the elementary schools in their districts. 93 per cent of the superintendents replied—all but those from four relatively small districts. All data from both sets of questionnaires were summarized and certain comparisons made using the chi square statistic.

A number of limitations should temper the conclusions drawn as a result of the findings below.

1. In 1968 an urban district could rightly be defined as one having a locally appointed superintendent. By 1971 this test for an urban district was no longer completely appropriate. All districts which were designated urban in 1968 were placed in the same category in 1971, although some moves toward or away from urbanity may have taken place.

2. As has been pointed out above, the four categories of staff utilization investigated are somewhat arbitrary when we recall the tremendous diversity of staff utilization practice.

3. The preferences for staff utilization patterns, solicited in 1968 from superintendents, principals and teachers, are recognized as weak indications of the direction which practice might take. The availability of adequate staff and facilities are only two factors which might cancel out the effect of attitude.

Findings

Staff Utilization in 1968.

The great majority of elementary teachers were teaching all or most subjects to a single group of children in 1968. Comparatively little special-

TABLE 1
PATTERNS OF STAFF UTILIZATION IN THE
ELEMENTARY SCHOOLS OF ALBERTA, 1968

Rural - Urban	Team		Semi-	Self-
	Teaching	Departmentalized	Departmentalized	Contained
Rural:				
I	*.2 (1)	.7 (3)	24.7 (112)	74.4 (338)
II	.8 (4)	7.1 (34)	40.5 (193)	51.5 (245)
Urban:				
I	3.4 (13)	.5 (2)	22.5 (86)	73.6 (282)
II	8.8 (33)	7.5 (28)	65.7 (245)	18.0 (67)
Total:				
I	1.7 (14)	.6 (5)	23.6 (198)	74.1 (620)
II	4.4 (37)	7.3 (62)	51.6 (438)	36.7 (312)
Grand Total:				
	3.0 (51)	4.0 (67)	37.7 (636)	55.3 (932)

* .2 percent (1 school) of all rural elementary schools having one or more primary classrooms uses team teaching. All figures in Table 1 conform to this format.

ization, either in the form of departmentalization or team teaching, was in evidence. There were, however, some interesting differences between primary and intermediate patterns of staff utilization: there was more team teaching, departmentalization and semi-departmentalization in the intermediate grades of Alberta, that is, relatively more specialization, and, relatively, about twice as many self-contained organizations in primary. Marked differences were found between the relative numbers of rural and urban patterns, with more team teaching in urban schools, more semi-departmentalized organizations in urban intermediate, and almost three times as many self-contained intermediate schools in rural (Table 1).

The preferences of teachers, principals, and superintendents in 1968 for the various patterns of staff utilization in the elementary schools of Alberta provided a dramatic contrast with the actual incidence reported above. While other factors, such as the availability of suitable staff and facilities, must exert powerful influences on the staff utilization patterns, it would seem likely that the preferences of key educators would be the most important factor determining educational practice.

The superintendents in Alberta in 1968 expressed a preference for team teaching in urban schools and for semi-departmentalized patterns in rural schools. Across the province, team teaching was preferred

TABLE 2
SUPERINTENDENTS' PRINCIPALS' AND TEACHERS' PREFERENCES
FOR PATTERNS OF STAFF UTILIZATION IN THE ELEMENTARY
SCHOOLS OF ALBERTA EXPRESSED IN PERCENTAGE OF FIRST
CHOICES, 1968

Position	Rural - Urban	Team Teaching	Depart-mentalized	Semi-Depart-mentalized	Self-Contained
	n.				
Superin-tendent:	Rural:				
	I (56)	*30.4	0	60.7	8.9
	II (57)	43.9	31.6	21.1	3.5
	Urban:				
	I (21)	33.3	0	47.6	19.0
	II (21)	52.4	33.3	14.2	0
Principal:	Total:				
	I (77)	31.2	0	57.1	11.7
	II (78)	46.2	32.1	19.2	2.6
	Rural	(29)	24.1	34.5	6.9
	Urban	(44)	56.8	9.1	0
	Total	(73)	43.8	19.2	2.7
Teacher:	Rural	(336)	17.0	19.3	41.1
	Urban	(512)	28.1	10.5	42.2
	Total	(848)	23.7	15.2	41.7

* 30.4 percent of rural superintendents ranked team teaching first as their preferred method of staff utilization for primary grades. All figures in Table 2 conform to this format.

for intermediate grades and semi-departmentalization for primary. Both rural and urban teachers preferred semi-departmental organizations. Team teaching received less support, and the self-contained classroom more support, from teachers than from either principals or superintendents (Table 2).

Staff Utilization in 1971.

As in 1968, the great majority of teachers taught all or most subjects to a single group of children: however even a casual comparison between the data for the two years would suggest a decided trend towards the sharing of at least some subjects. There was, then, much more specialization than in 1968.

The sharp contrasts between primary and intermediate patterns of staff utilization, noted in 1968, were still in evidence in 1971: there was more team teaching, departmentalization, and semi-departmentalization in intermediate grades and more self-contained organizations in primary. However, specialization had increased in primary at a faster rate during the three year period.

Rural-urban differences were still notable in 1971. Even greater differences were found in the incidence of team teaching and of self-contained classrooms in rural and urban schools (Table 3).

TABLE 3
PATTERNS OF STAFF UTILIZATION IN THE ELEMENTARY
SCHOOLS OF ALBERTA, 1971

Rural - Urban		Team Teaching	Departmentalized	Semi- Departmentalized	Self- Contained
Rural:					
	I	*3.0 (14)	1.7 (8)	39.7 (186)	55.6 (260)
	II	4.0 (18)	12.3 (55)	53.3 (239)	30.4 (136)
Urban:					
	I	13.2 (49)	2.2 (8)	40.8 (151)	43.8 (162)
	II	18.0 (65)	8.0 (29)	65.7 (238)	8.3 (30)
Total:					
	I	7.5 (63)	1.9 (16)	40.2 (337)	50.4 (422)
	II	10.2 (83)	10.4 (84)	58.9 (477)	20.5 (166)
Grand Total:		8.9 (146)	6.1 (100)	49.4 (814)	35.7 (588)

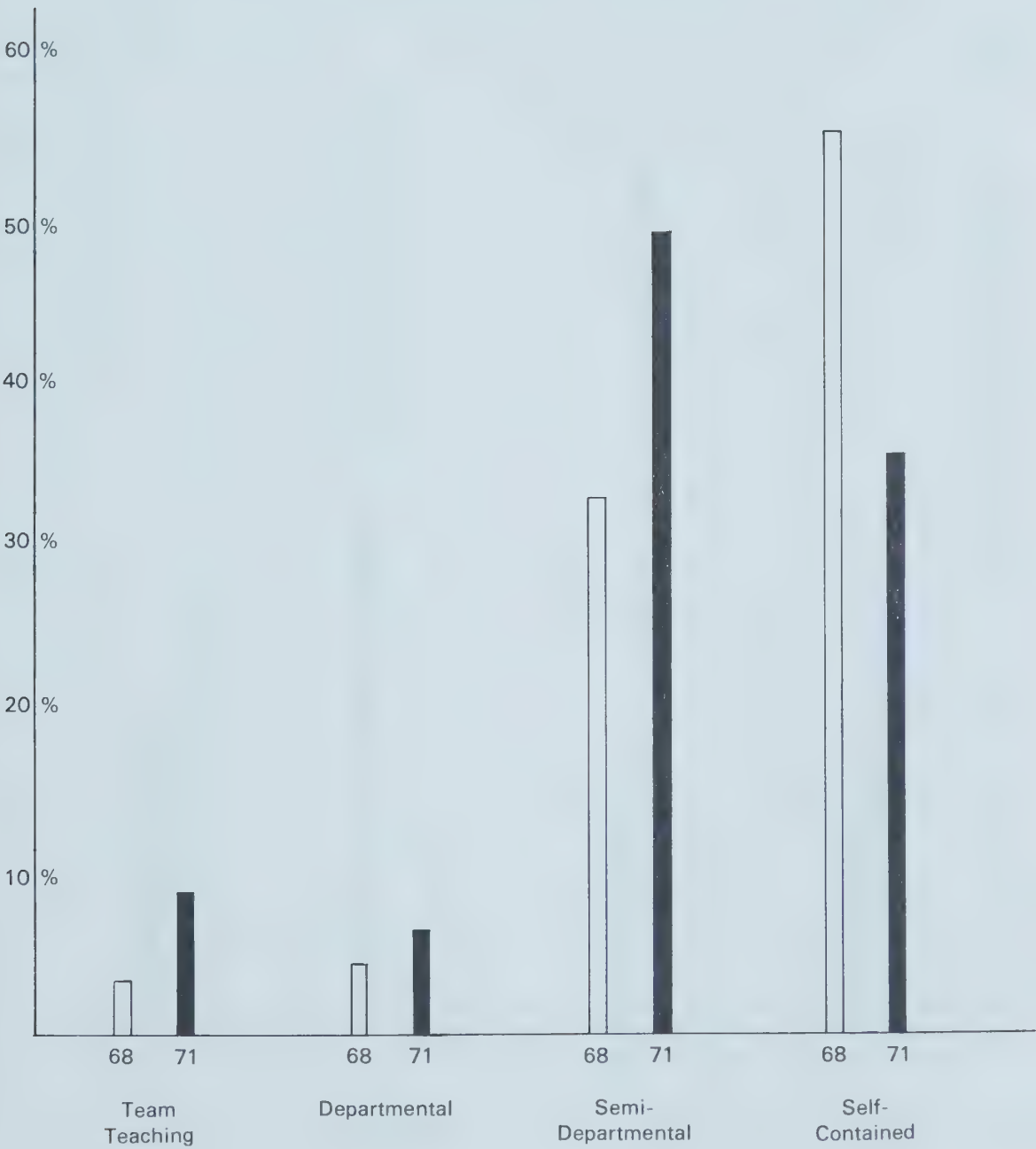
* 3.0 percent (14 schools) of all rural elementary schools having primary classrooms use team teaching. All figures in Table 3 conform to this format.

1968 vs. 1971

A significant difference was found when the 1968 results were compared with those of 1971 ($\chi^2 = 151.62$, $p > .001$). The most dramatic change was in the decrease in the incidence of the self-contained classroom. This is attributable to an increase in specialization in the form

of semi-departmentalization and, to a lesser degree, in team teaching and departmentalization (Figure 1).

FIGURE 1
1968 Compared with 1971 Patterns
of Staff Utilization

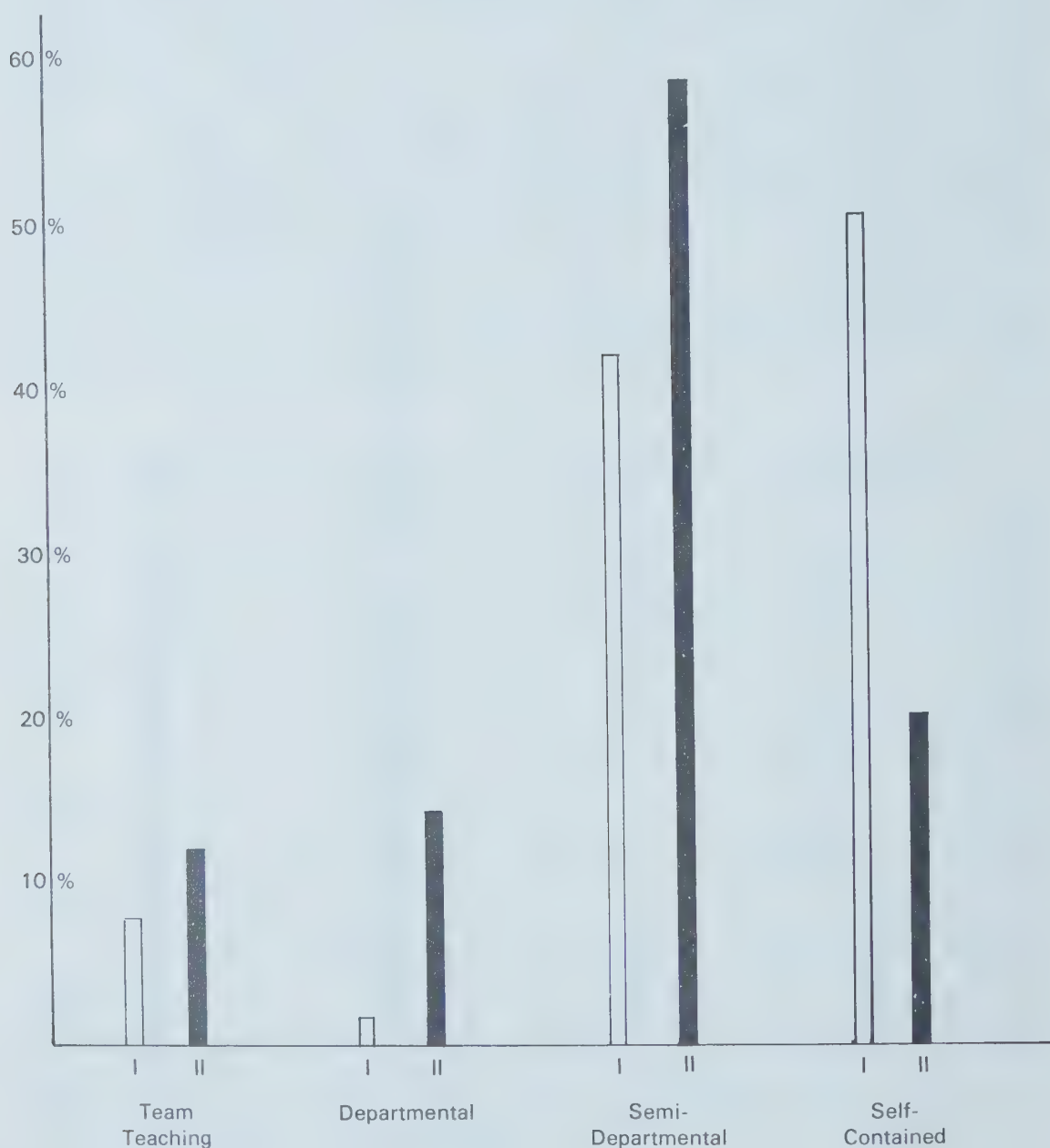


1971 Results: Primary vs. Intermediate

A significant difference was found between primary and intermediate patterns of staff utilization in 1971. ($\chi^2 = 184.09$, $p < .001$). About two and one-half times as many self-contained organizations were found in

primary compared with intermediate. More team teaching, semi-departmental and departmental organizations were found in intermediate grades (Figure 2).

FIGURE 2
Primary Compared with
Intermediate Patterns of
Staff Utilization, 1971

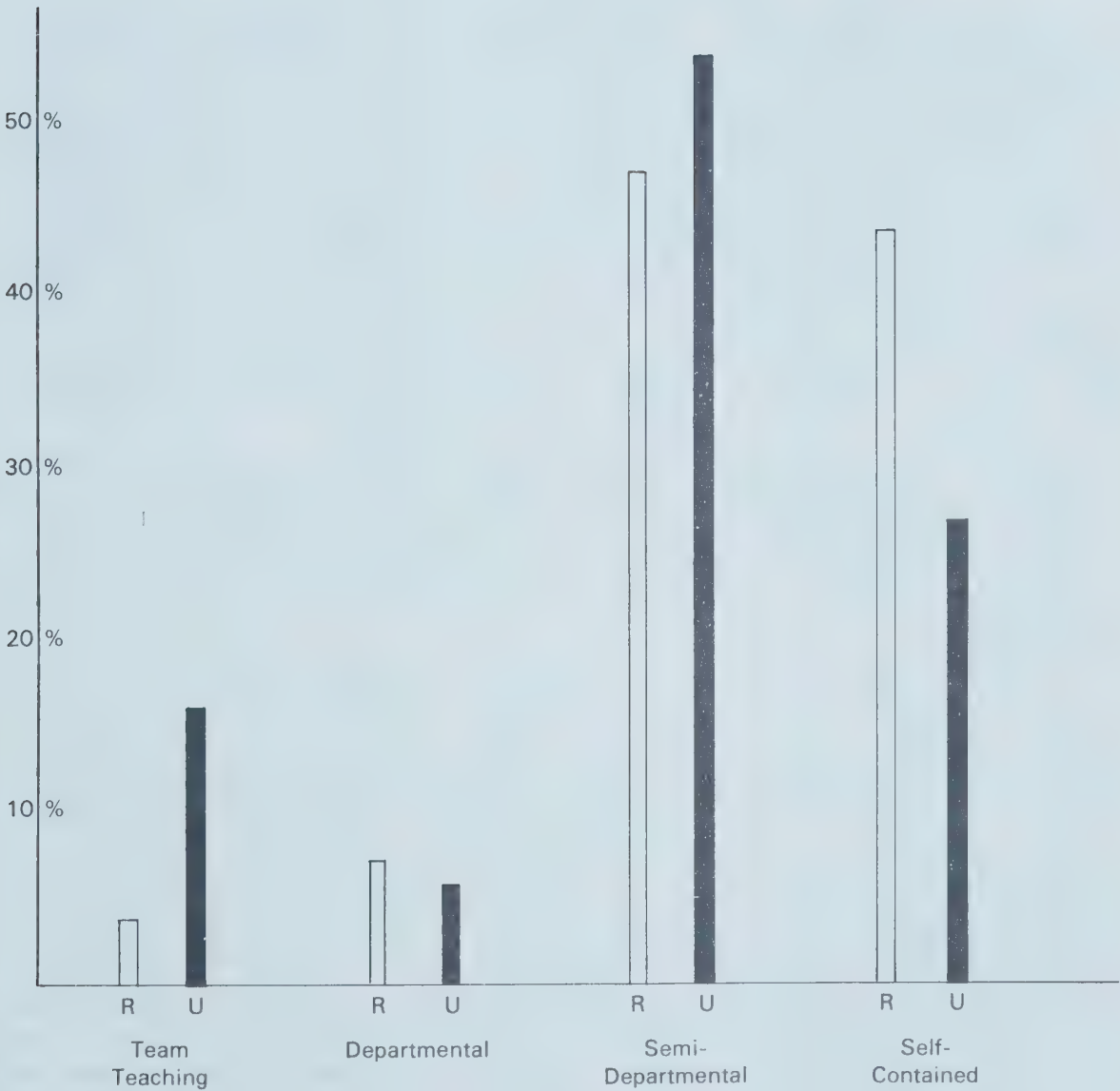


1971 Results: Rural vs. Urban

Patterns of staff utilization in elementary schools in 1971 were found to be significantly different depending upon whether the schools were found in rural or urban Alberta ($\chi^2 = 105.95$, $p < .001$). The greatest

differences existed in self-contained classrooms and in team teaching: many more self-contained classrooms were found in rural Alberta and more team teaching in urban Alberta (Figure 3).

FIGURE 3
Rural Compared with Urban Patterns
of Staff Utilization



1971 Results: Patterns by Size of School

The patterns of staff utilization were found to be significantly related to the size of the school. The small schools whether rural or urban, primary or intermediate, tended to have a higher than average incidence

of self-contained classrooms. The larger schools tended to use more staff specialization including more team teaching.

1971 Results: Patterns by Size of District

Except for rural primary, the incidence of patterns of staff utilization was found to be significantly related to the size of the school district as measured by the number of instructional units in the district. (The primary division of a school, for example, might be described as an instructional unit for purposes of the present survey). The largest districts tended to have a higher than expected frequency of team teaching.

Conclusions and Implications

1. The findings of this study indicate a major change in the elementary schools of Alberta over a relatively short period of time. The implications for teacher education are many. For example, about one out of every ten of the elementary schools in Alberta was using team teaching in 1971. The curriculum of teacher education institutions, with the current emphasis on the self-contained classroom, would, consequently, seem to be in need of change. Optional pre-service courses, providing for the development of team teaching skills, might be offered. In addition, since pre-service education is over for many teachers who are now a part of new organizational plans, the responsibilities of faculties of education must be extended to match the professional life of the teacher—from admission to teacher education until retirement.

2. The great speed of the change reported here is disturbing. Ideally, change is a rather deliberate process, based on the evaluation of existing practices, careful planning and provision for self study and inservice programs. The process described is time consuming and does not permit the amount of pendulum-swinging which might be one interpretation of the present data.

3. Substantial differences in patterns of staff utilization between certain groups of schools were noted. Rural and urban schools, primary and intermediate grades, small elementary and large elementary schools, and small school districts and large school districts were all significantly different in terms of staff utilization patterns. The reasons for the differences are only partly known but all differences noted would probably suggest differences in the amount or in the nature of the pre-service and in-service education to be made available. For example, the rural-urban differences might suggest that change is easier in urban centres where more new schools are built each year. Without presuming to place value on the various patterns of staff utilization, the impetus for change has been felt to a much lesser degree in rural areas, at least in that at least in the area studied in the present study. If this apparent slowness of change is characteristic of rural areas, perhaps the pre-service and/or in-service education for rural areas needs to be different from that provided for urban areas.

4. The actual changes in patterns of staff utilization are interesting, but the reasons for these changes are even more important. For example, do subject-centred individual differences suggest that greater teacher

specialization is perhaps taking place at the expense of the child as a person? Research to identify reasons for change is badly needed: this study suggests the need for major research studies beyond the scope of individuals.

References

- Anderson, Robert H. Some types of cooperative teaching in current use, *The National Elementary Principal*, 1965, 44, January.
- Blench, W. A. A study of the teaching of science in the elementary schools of Alberta. Unpublished master's thesis, The University of Alberta, 1967.
- Coultas, H. D. A study of organization for music education in the elementary schools of Alberta. Unpublished master's thesis, The University of Alberta, 1965.
- MacKay, D. A. and Ward, K. L. Team-teaching in western Canada, *The Canadian Administrator*, 1967, 7, December.
- Polos, N. C. *The dynamics of team teaching*. Dubuque, Iowa: Wm. C. Brown, 1965.
- Ratsoy, E. W. Tandem teaching: a new approach, *The C S A Bulletin*, 1968, 7, June.
- Shaplin, J. T. and Olds, H. F. (Eds.) *Team teaching*. New York: Harper and Row, 1964.

FARHAT GHAEM-MAGHAMI

University of Toronto

Political Climates, Political Knowledge and Political Preferences of University Students

This article examines the impact of a number of socio-political variables and the political environment — the national political opinion climate — on the party choice of university students using data from 1,006 questionnaires completed by first year University of New Brunswick students.

Political attitudes and political socialization of the student, have received increasing attention in recent years. The lowering of the voting age to 18 has given young voters an opportunity to influence the outcome of any national election. The continuity of political party choice across generations is well documented. Yet, the choice of a political party, as Levin (1961) points out, is influenced by "the political climate of opinion in various subsystems of which the adolescent is a member and contributes to his political socialization and his choice of a political party" (Levin, 1961, p. 596).

This article reports research which considers the influence of some background variables on the political party preferences of university students in the province of New Brunswick.

Levin, employing data collected from students in ten northern Illinois high schools, found the best predictor of the students' party choice to be the party choice of their parents (Levin, 1961, p. 596). Johnson (1972), in a similar work in six rural Kentucky counties, reports that "the overwhelming factor influencing youth's political party preference is the preference of parents" (Johnson, 1972, p. 51). However, there are other factors, besides parental influence, which affect the choice of political party. On the one hand, there are social structural variables such as sex, socioeconomic status (SES) and religion which influence the voting patterns. For instance, Blake reports that in Canada "Roman Catholics show disproportionate support for the Liberals" (Blake, 1972, p. 71). Moreover, on a regional basis, he demonstrates that there is less Liberal support, amongst the white-collar group in the Maritimes and more in British Columbia (Blake, 1972, p. 71). On the other hand, there are

psychological feelings and knowledge about political matters which can also influence the choice of a political party. Political leaders might be wondering as to how an individual's self-definition as a liberal or a conservative affects his voting patterns. After all, a student who chooses to define himself as a liberal will not be likely to contradict himself and vote for the Progressive Conservative party or vice-versa. Another question is: How do the alienated and relatively unalienated youth differ in their voting patterns? What is the impact of a person's level of political knowledge on his political choice?

The aim of this research is to examine the effects of the above-mentioned variables on the voting patterns of students, but, also to test the contention that almost nine years of Liberal government has produced some variation in favor of the Liberals.

Method

Data were collected from first year students in the University of New Brunswick. Questionnaires were administered in the early fall of 1971, after nine years of a federal Liberal government. Consequently, these students who were entering the adult political world, had spent the years from ten to eighteen in a national political opinion climate (operationally defined by the party in power at the national level) that was Liberal.

Respondents' political preferences were ascertained by asking them to name the party they would vote for in an election and also by their political self-definition. Father's political preferences were also reported by respondents.

Measurement of alienation, political knowledge, and socioeconomic status (SES) was accomplished by appropriate scales which are reported elsewhere by the author (Ghaem-Maghami, 1972). Briefly, following Seeman (1966), alienation is measured by a ten item powerlessness scale: scores were dichotomized at the mean. Those who scored above the mean were classified as alienated. The political knowledge scale was operationalized by a twenty-five item information test which

TABLE 1
POLITICAL PARTY CHOICE BY SEX AMONG UNIVERSITY STUDENTS
(in percentages)

Party Choice	Male (N=650)	Female (N=356)	Total (N=1006)
Progressive Conservative	22	17	20
Liberal	38	48	42
New Democratic Party	16	10	14
Undecided	23	25	24
Total Per Cent	100	100	100

$\chi^2 = 12.71$; 3 df, $p < .01$.
Cramer's V = .11.

required the student to identify political authorities and organizations in the Canadian and international scene as in the following examples:

- 1. Who is the Finance Minister of Canada?
- 2. What do the initials NATO stand for?

Students unable to identify more than five items were categorized as having low political knowledge. SES was measured by a composite index which combines education, income, and occupation of parents of the respondents.

Findings

The data show that the Liberal party is much more favored among these students than either the Progressive Conservatives or the New Democratic Party (N.D.P.). As is shown in Table 1, 42 per cent of students indicated that they would vote for the Liberals in an election, as compared to 20 per cent for the Conservatives and 14 per cent for the N.D.P. However, the Liberal party is favored more than three to one by female students, against the Conservatives (48% vs. 17%), and almost five to one against the N.D.P. (48% vs. 10%). It should be noted, however, that twenty-five per cent of students are undecided about their political party choice in a future election.

Although it has been found in previous studies that Roman Catholics show disproportionate support for the Liberals, Blake (1972, p. 70) reports a high level of Conservative support amongst Catholics in the Prairie Provinces and in the Maritimes. To examine this trend in our data we controlled for religion. Table 2 demonstrates that, at least among

TABLE 2
POLITICAL PARTY CHOICE AND RELIGION OF THE STUDENTS
(in percentages)

Party Choice	Catholics and Others (N=342)	Protestant (N=664)	Total (N=1006)
Progressive Conservative	16	22	21
Liberal	44	40	42
New Democratic Party	16	13	14
Undecided	24	25	24
Total Per Cent	100	100	100

$\chi^2 = 6.66; 3 \text{ df}, p < .10.$

students, this trend has been reversed in favor of the Liberals. Only 16 per cent of Catholics say they will vote Conservative as compared with 44 per cent who would vote for the Liberals. Moreover, Conservatives can count only on 22 per cent of the Protestant vote among these students—almost half that of the Liberals (40%). Blake reports that “the white-collar group is also the only class group which appears to differ by

region, with less Liberal support in the Maritimes and more in British Columbia" (Blake, 1972, p. 71). To see if the national political climate has affected this trend we controlled for various measures of socio-economic status. Table 3 presents the party choice and father's occupation

TABLE 3
POLITICAL PARTY CHOICE OF STUDENTS ACCORDING TO THEIR
FATHERS' OCCUPATIONS
(in percentages)

Party Choice	White-collar (N=517)	Blue-collar (N=489)	Total (N=1006)
Progressive Conservative	17	24	20
Liberal	45	38	42
New Democratic Party	14	14	14
Undecided	24	24	24
Total Per Cent	100	100	100

$\chi^2 = 8.63$; 3df, $p < .05$.

Cramer's $V = .10$.

among students. However, it should be noted that the cross-tabulations with other dichotomized measures of social class such as the SES scale, father's income and father's education produced almost an identical result. These results show that there exists a slight tendency (7%) for the high SES group to vote for the Liberal party and for the low SES to vote for the Conservative party (also 7%). It is shown in Table 3 that 38 per cent of students with blue-collar fathers as compared to 45 per cent of those of white-collar origin would vote for the Liberals. On the contrary, 24 per cent of those whose fathers had a blue-collar job would vote Conservative, as compared to 17 per cent of those whose father had a white-collar job. The tendency on the part of students with white-collar fathers to vote Liberal can be interpreted as being due either to the general liberalization of students, or to the influence of the national political climate under the Liberal party. None of the measures of social class made any difference to the 14 percent support which is expressed for the N.D.P. This testifies to the fact that the N.D.P. has a low-level of appeal for the working, as well as the middle class, students.

The question to be asked is how these students differ from their parents in their choice of political party? Further, how does the national political opinion climate, generated by the party in power, affect the student's political party choice?

The data substantiate many earlier findings that the overwhelming factor influencing youth's political party choice is the voting choice of parents. As is shown in Table 4, the Liberals have a higher "retention rate" (i.e. the percentage of students identifying with the party of their father) than the Conservatives (63% versus 45%).

TABLE 4
FATHERS' REPORTED PARTY CHOICE AND PARTY CHOICE OF STUDENTS

Father's Party Choice	Percent Students Choosing Same Party as Father	N
Progressive Conservative	45	272
Liberal	63	414
New Democratic Party	55	22
Undecided	39	298

If, as has been suggested, a new generation tends to affiliate with the party in power during its political socialization, we should find not only an increase in the rate of retention of this new generation for the Liberals; but, also an increase in the rate of conversion amongst those of Conservative parents, to the Liberals.

TABLE 5
THE OVERALL COMPARISON OF FATHER'S REPORTED PARTY CHOICE
AND PARTY CHOICE OF STUDENTS

Student's Party Choice	Father's Party Choice			
	Conservative (N=272)	Liberal (N=414)	N.D.P. (N=22)	Don't Know (N=298)
Progressive Conservative	45	7	—	17
Liberal	21	63	5	33
New Democratic Party	13	14	54	11
Undecided	21	16	41	39
Total Per Cent	100	100	100	100

$\chi^2 = 279.83$; 9 df, $p < .001$.

Cramer's $V = .30$.

Table 5 presents the overall comparison of students' political party preference and that of their parents. Of significance are findings that only seven per cent of children of Liberal parents would vote for Conservatives as compared to twenty-two percent of youths from Conservative parents who would vote Liberals. The findings seem to substantiate Levin's contention about the influence of the national political climate on the choice of political party.

One qualification should be made, however, before making such a generalization: we contend that the student's political self-definition intervenes in the relationship between parental political affiliation and the student's voting preferences. That is to say, the difference in the reten-

tion rate might be an artifact of the differential political self-definitions among these students.

TABLE 6

FATHER'S REPORTED PARTY CHOICE AND PARTY CHOICE OF STUDENT,
BY STUDENT'S POLITICAL SELF-DEFINITION†
(in percentages)

STUDENT'S POLITICAL SELF-DEFINITION					
Student's Party Choice	Conservative*		Liberal**		Total (N=510)
	Father's Party Choice Conservative (N=98)	Liberal (N=49)	Father's Party Choice Conservative (N=93)	Liberal (N=270)	
Progressive					
Conservative	81	29	34	5	27
Liberal	13	49	43	80	58
New Democratic Party	6	22	23	15	15
Total Per Cent	100	100	100	100	100

* $\chi^2 = 38.07$; 2 df, $p < .001$. ** $\chi^2 = 64.99$; 2 df, $p < .001$.
Cramer's V = .51. Cramer's V = .42.

† Note: For computational purposes students who were undecided on their choice of a political party and the students whose fathers voted for the N.D.P. were excluded.

Table 6 shows that, when the students' political self-definition is controlled, the Conservative and the Liberal party have a similar retention rate (80% versus 81%). According to our data, however, the Liberal party shows a higher overall rate of retention than the Progressive Conservative party does. Although these data involve only one community, they are sufficient to call in question the hypothesised influence of national political climate, at least among students in Maritime Canada.

For the politicians, as well as academicians, two questions remain. First, how is the "alienated voter" likely to vote in an election? and second, where would the votes of those who "know-nothing" about politics go? The "radicals" might argue that the alienated voters will choose the N.D.P. since it represents an "alternative" to the status quo. In the same vein, "critics" might argue that the bulk of the know-nothing voters "who want to jump on the bandwagon" tend to swing the pendulum of an election in favor of the Liberal party now in power.

To explore these questions, we utilized our alienation and political knowledge scales. If the critics' opinions were true we should find the bulk of alienated students choosing the N.D.P. and the bulk of the "know-nothing" voters choosing the Liberals. Table 8 shows the voting patterns of students, controlling for alienation and political knowledge

TABLE 7
POLITICAL PARTY CHOICE WITH ALIENATION AND POLITICAL
KNOWLEDGE AMONG UNIVERSITY STUDENTS
(in percentages)

Party Choice	Alienated*		Unalienated**		Total (N=1006)
	Political Knowledge High (N=200)	Low (N=310)	Political Knowledge High (N=279)	Low (N=217)	
Progressive Conservative	19	27	14	20	20
Liberal	37	42	46	40	42
New Democratic Party	23	7	18	10	14
Undecided	21	24	22	30	24
Total Per Cent	100	100	100	100	100

* $\chi^2 = 26.48$; 3 df, $p < .001$.
Cramer's V = .23.

** $\chi^2 = 10.80$; 3 df, $p < .02$
Cramer's V = .15.

simultaneously. In Table 7 we can detect only a slight tendency in favor of the above hypothesis. There are various ways of interpreting Table 7; but, briefly, it appears that the Liberal party tends to get its greater support (46%) from the unalienated and high political-knowledge group. The N.D.P. tends to be supported more by the alienated and high political knowledge group (23%). The Conservatives, on the other hand, tend to find greatest support amongst the alienated low political knowledge group (27%) and least support amongst the unalienated who have high political knowledge. The differences, however, are not so great to allow us to generalize beyond our data.

To the extent that these students identify the Liberal party as the party of the middle class, aspirations for social mobility might account for the greater loss by the Conservatives. It is also of prime importance to include in the operational definition of national political opinion climate, besides the party in power nationally, the general political mood of a nation in an historical epoch. Many people, especially students and the young, would like to define themselves as liberals and this self-definition would affect their voting behavior on the national level. Although there is now a federal Liberal government, it is accompanied by a national political mood in a direction favorable to the Liberal party.

Discussion

This research has generally supported the effect of national political climate defined by the party in power at the national level and the general political mood of the voter. Amongst all the variables considered, student political self-definition and parental political affiliation are the best predictors of his party choice.

In considering these data, two purely speculative notions are discussed: first, assuming that the greater retention rate of the Liberals is a consequence of the students' aspiration for social mobility, such a process might be necessary to maintain a clear majority. It is possible that upward mobility leads children of the working class to move to the middle class party thereby assuring the majority necessary for an effective federal government.

Second, the effects of national political opinion climate also seem to make for stability at the federal level while permitting the continuation of differential party choice at the provincial level. When a majority party at the national level retains a high proportion, and the minority parties a lower proportion, of the next voting generation, the results will be a relatively stable alignment. Moreover, it appears that, during the time when most of the consumers can buy what they want (with cash or credit) and the leisure-oriented youth can get some pocket-money (through programs such as Opportunities for Youth), the rhetoric of the federal minority parties about "full-employment," inflation, and recession will only appeal to a small audience among these youth.

References

- Blake, Donald E. The measurement of regionalism in Canadian voting patterns, *Canadian Journal of Political Science*, 1952, 5, 53-81.
- Ghaem-Maghani, Farhat. *Powerlessness and political knowledge: a study of the political orientations of freshmen at the University of New Brunswick*. Unpublished M.A. Thesis, the University of New Brunswick, 1972.
- Johnson, M. R. Political climates and party choice of high school youth, *Public Opinion Quarterly*, 1972, 36, 48-55.
- Levin, M. Social climates and political socialization, *Public Opinion Quarterly*, 1961, 25, 596-606.
- Seeman, Melvin. Alienation, membership, and political knowledge: a comparative study, *Public Opinion Quarterly*, 1966, 30, 353-367.

CLIFFORD D. PENNOCK

University of British Columbia

Selecting Instructional Materials to Meet Students' Reading Performance Level

Three major approaches to the selection of appropriate level reading instruction materials are described and evaluated. Strengths of each procedure are noted, as well as sources of weakness in test construction, administration, and the interpretation of scores. Specific suggestions are made with a view to improving reliability and validity. Use of the cloze procedure in selecting reading materials is recommended, since cloze appears to combine the advantages of both standardized and informal procedures.

Programs designed to individualize instruction in social studies, mathematics, science, and English usually require going beyond the single text to a variety of sources. Texts, tradebooks and other materials often appear suitable in topic content and interest, but more information on their readability is often needed. Specifically, one may wish to ask these questions: Could the selection under consideration be read with understanding by any one student in the class? By a small group? By most students? By all? The number of copies purchased, the use made of the book, and the need for alternative materials will depend to some extent upon the answers to the previous questions.

Three major approaches to selecting appropriate level reading materials will be discussed:

1. Readability—Standardized Test Procedure
2. The Informal Reading Inventory
3. The Cloze Procedure

Readability—Standardized Test Procedure

Determining Readability

The readability of a text is sometimes provided by the publisher in the form of a recommendation that the material is suited to a school grade; for trade books, an age-interest-suitability is more frequently pro-

vided; occasionally a readability formula is referred to, and an estimate of readability provided, representing, one would hope, the mean of several sample indices from the reading materials in question. Quoting such an index or mean readability of a book as *the* readability, however, may too often lead the prospective buyer into the false sense of security, that no parts or few parts of the book are more difficult than the (average) readability quoted. Much to be preferred would be the provision, also, of the *range* of difficulties within the book, as well as a statement of the number of samples taken to provide the data.

Estimating the readability of a book is only a first step. One must also assess the varying performance levels at which the book's potential readers function.

Assessing Performance Through Standardized Tests

Harris (1965) suggests that standardized reading tests can be used as an "aid in the selection of reading materials of appropriate difficulty," with grade scores "tending to show the instructional level" of the student to which an appropriate match can be approximated with materials of ascertained readability. Admittedly this is a process of approximation which includes error components from several sources—the less than perfect stability, homogeneity and validity of any test, as well as the less than perfect validity of the readability formula used, and any inaccuracy in its application. Substituting use of an informal inventory of unproven reliability or validity would appear unlikely to reduce error, although it might be supported in other ways, to be discussed later.

Difficulties

Heilman (1967) contends that students, in responding to standardized (silent) group reading tests, do not require complete understanding to get correct answers. This statement can be supported by advancing the possibility of random correct scores, and also of obtaining a correct response on every item through only partial knowledge, although such a distribution with no more than partial comprehension of every item would seem rather unusual, to say the least.

Harris (1965) suggests that poor readers tend to guess more than good readers and that their scores may then be overestimated on standardized tests.

If in contrast, one takes Goodman's (1970) position that reading is a "psycholinguistic guessing game," the point might be taken that the good readers guess more effectively than the poor readers. It would certainly seem plausible that good readers would make more effective use of minimal knowledge than poor readers. On the other hand, it is possible for a score to be inflated by correct answers obtained by random choice. Such effects can be minimized by built-in penalties for guessing. Exact-word completion response requirements, of course, foreclose this abuse, but appear to necessitate hand-scoring.

Recent evidence presented by Mitchell (1967) Weaver and Bickley (1967) and Bradley (1970), support the contention that reading compre-

hension tests are measuring general knowledge as well as a variety of influences connected to the test *questions* in addition to the assumed task—the ability to answer items from cues in the reading passage.

While it might be contended that extent of general knowledge as well as the ability to gain information from test questions are, perhaps, indicators of reading comprehension, the phenomena would appear to inject error into such tests in that they are not measuring in toto what is assumed—comprehension of a set of passages. A logical remedy would seem to be the removal of multiple choice formats in reading tests, with information gain indicated through the cloze procedure.

Parenthetically it might be suggested, in jest, that the findings cited above imply, also, the possibility of conserving considerable test form space and paper, through eliminating the passages and marking responses on a set of multiple choice statements in isolation.

It would seem quite possible, however, that the question-answering process approaches quite closely what can be termed verbal ability or verbal intelligence. If this be so, then some error is thereby introduced, although the net effect in predicting comprehension for specific materials would not appear to preclude use of such tests for that purpose, due to the known high correlation between verbal intelligence and reading comprehension.

Improving Test Selection

Chall (1970) has suggested a practical maneuver to increase the effectiveness of standardized test use. Since standardized tests are mainly designed for a spread of a few grades, they tend to give an inaccurate estimate of the high achiever through ceiling effects, and to be unreliable for the poorest readers through too few items at his level. Chall suggests that test selection for students who tend to score at the extremes should be made on the basis of their estimated reading level rather than on grade placement.

A grade six youngster, for example, who might be expected by his teacher to score at the grade nine or ten level on test forms addressed to the intermediate grades, would be given, instead, a test form designed for high school students. Similarly a student in the fifth grade having considerable difficulty in reading would be administered a primary grade form, probably followed by individual testing as well.

The Informal Reading Inventory

Face Validity

Asking a student to read selections aloud and at sight would appear on the surface, perhaps, as a very direct, practical, and appropriate method of matching the difficulty of a book with a student's reading performance level. It has one appealing advantage over the use of standardized tests, group or individual, oral or silent: no error is introduced in making a readability estimate nor in matching it with test scores, since the testing is done with the materials for instruction, and no readability estimate needs to be made.

Criteria Validity Questioned

Powell (1971) has presented evidence based on two comprehensive empirical investigations carried out at the University of Illinois (Powell, 1969; Dunkeld, 1970), supporting his contention that the conventional informal word recognition criterion (95%), for the instructional level, is unduly stringent.

Difficulty of Administration

Teachers, even after intensive training, appear unable to gain sufficient proficiency in test administration, in recognizing oral reading errors and in interpretation of results (Emans, 1965; Ladd, 1961).

Such lack of uniformity in test administration, scoring and interpretation of results, would appear likely to introduce considerably more error variance into informal testing, in contrast to the use of standardized tests where a great effort is usually made to make administration, scoring, and interpretation clear-cut and unambiguous.

Unrealistic Construct of Reading Performances

A basic question is that of validity: Can one select materials to be read *silently* on the basis of *oral* reading performance? Probably not, beyond grade one. Spache (1969), for example, summarizes considerable data to support the concept of a trend toward a smaller relationship between oral and silent reading ability as students mature, with the prediction of silent reading from oral reading errors approximately 36% better than chance in the elementary grades and 25% better than chance in high school. Similarly a recent study by Becker (1970), with fourth grade pupils, included the findings that those students with the lowest comprehension scores made the fewest oral reading errors, while the best comprehenders made the most errors in word calling.

One may conclude that scoring word recognition errors in a student's oral reading does have a purpose—it provides a record of the kind of word recognition errors made by this student in a specific but somewhat artificial (sight) reading situation; a record possibly of value in diagnosis and prescription of word attack instruction.

Scoring word attack errors does not appear to be a useful practice, however, if the purpose of the reading is to discover whether the material is at a suitable comprehension level, i.e. whether, for example, a particular social studies book is appropriate for this youngster's reading performance level. If oral reading is to be used, the criterion for instruction should be based on the percentage of comprehension questions answered correctly, generally 70% (Powell, 1969), or 75% (Johnson and Kress, 1965).

The situation is still somewhat artificial, in that one is asking a student to answer questions after reading aloud, when most of the reading he will later do on the materials in question will be silent reading.

Advantages of an Oral Format—Informal or Standardized

There are, however, two advantages in an oral format which should bring it into use where advantageous: (a) the teacher listens to the student

reading aloud and knows the student is attacking the words rather than indulging to a variable extent in off-task behavior; (b) the teacher asks the student the comprehension questions eliminating the possibility that the student cannot, or does not read each question. It is possible, of course, to ask a student to read a selection silently and then to respond to oral questions, thus eliminating the questionable practice of evaluating oral reading for silent reading purposes. Greater control is also thus maintained over the question-answer process.

Comprehension Questions—Guidelines for Construction

It remains necessary to compose a variety of questions for each passage, in sufficient number to ensure a modicum of reliability, and of sufficient quality to ensure measurement of what one wishes to define as comprehension. Questions must also be formulated in such a way as to preclude the possibility that the student may obtain the correct responses through the questions themselves or through previous experience rather than from the passages in question.

Care must be taken that questions posed on selections can be understood either through reading or listening, depending upon the input medium used, since it is assessment of comprehension of the passage read that is our task, rather than the evaluation of the student's ability to read the questions.

Evaluation of Oral Reading—Uses and Limitations

Identification of appropriate reading material through the observation of oral reading would, then, appear beset by many difficulties, making its use most appropriate as a diagnostic and supplementary evaluation procedure.

The Cloze Procedure

A variety of studies indicate the validity, reliability and utility of the cloze procedure in evaluating a student's comprehension of reading selections. Extensive description and criticism of this research can be found in Bormuth (1968), Greene (1968), and Rankin (1964).

Bormuth (1968) suggests the following procedures as most useful, "when validity, economy and convenience are considered simultaneously":. Every fifth word of the passages under consideration should be deleted, to be filled in by the student as he reads. This deletion ratio provides the maximum number of items for a passage ensuring reliability, while at the same time keeping the test to a minimum length (250 words is recommended). Answers are scored correct only when they match the deleted words, ignoring minor misspellings. If the student answers are correct on 40% or more of the deletions, the materials can be considered of appropriate difficulty for teacher directed instruction. When the student scores 60% or more, the material can be considered suitable for his independent work or study (Bormuth, 1967, 1968; Rankin and Culhane, 1969; and Lockerbie, 1969).

It is recommended (Bormuth, 1968) that a book to be evaluated through

cloze should first be sampled by administering cloze tests from six to twelve random samples of approximately 250 words, retaining for use that test having a mean closest to the mean of all tests. It seems unlikely that many teachers would be willing to engage in test production and administration to this extent, for the purpose of assessing the difficulty of one book with students, although it is possible such a procedure might be followed by a reading coordinator or similar person, to develop suitable school-wide tests for texts in common use. Copies could then be available in each school for duplicating and use by teachers.

A less time-consuming sampling process would be the following—six to twelve, 250 word samples might be selected randomly with two or more from the first third, the middle third, and the last third of the book. Readability of each sample could be calculated by some direct procedure, for example, Fry's (1968) readability graph. A 50 item cloze test might then be constructed from that sample closest to the mean readability of the samples examined. A student successful in 40% or more of his responses—20 answers correct, can be considered to be at the instructional level in this material. A student scoring 60% or more—30 or more correct responses can be considered at the independent level in this material. When a cloze test is administered to a group of children, similar safeguards should be taken as with any other group test. Particular care should be taken that students understand the directions, and that they are "on task" in attempting to replicate the author's message. Individual testing is to be advocated with students who for any reason would be handicapped on the test in a group situation. Re-testing on an individual basis and/or cross-checking by means of an informal inventory approach based on the same passages is indicated where individuals scores appear suspect.

Summary

Standardized test grade scores can be considered a rough estimate of a youngster's reading performance. Reading material can be approximated to this level through comparison of readability indicated, and grade scores, although this process combines the error inherent in any test with the error component of the readability formula and its application.

Informal (oral) evaluation procedures, beset by problems of scoring and interpretation, undetermined reliability, and an artificial construct of reading performance (oral sight reading), would appear least suited to identification of a silent reading performance level.

The Cloze Procedure appears to combine the advantages of both standardized and informal test procedures—the reliability, validity and much of the scoring ease of the former, with the pertinence and relevance of the latter, in that passages are constructed from instructional materials actually to be used by students.

Cloze would appear a valid and reliable tool to ascertain reading comprehension of instructional materials of most students beyond the first grade. Use of the Cloze Procedure is also considerably more convenient and considerably less time-consuming than use of the Informal Reading Inventory.

References

- Becker, M. *A Description of Observed Oral Reading Phenomena of Grade Four Children*. Unpublished Master's thesis, The University of Alberta, 1970.
- Bormuth, J. R. Comparable Cloze and Multiple Choice Comprehension Test Scores. *Journal of Reading*, 1967, 10, 291-299.
- Bormuth, J. R. The Cloze Readability Procedure. *Elementary English*, 1968, 45, 429-436.
- Bradley, M. Effects on Reading Tests of Deletions of Selected Grammatical Categories. In G. B. Schick and M. May (Eds.), *Nineteenth Yearbook of the National Reading Conference*, 19, 1970, 87-93.
- Chall, J. Interpretation of the Results of Standardized Reading Tests. In R. Farr (ed.), *Measurement and Evaluation of Reading*. New York: Harcourt, Brace and World, 1970, 51-59.
- Dunkeld, C. *The Validation of the Informal Reading Inventory for the Designation of Instructional Reading Levels: A Study of the Relationships Between Children's Gains in Reading Achievement and the Difficulty of Instructional Materials*. Unpublished doctoral dissertation, University of Illinois, 1970.
- Emans, R. Teacher Evaluations of Reading Skills and Individualized Reading. *Elementary English*, 1965, 42, 258.
- Fry, E. A Readability Formula that Saves Time. *Journal of Reading*, 1968, 11, 513-516, 575-578.
- Greene, F. P. Cloze Symposium. In G. B. Schick and M. May (Eds.), *Seventeenth Yearbook of the National Reading Conference*, 17, 1968, 110-131.
- Goodman, K., Reading: A Psycholinguistic Guessing Game. In H. Singer and R. Ruddell (Eds.), *Theoretical Models and Processes of Reading*. Newark: International Reading Association, 1970, 259-272.
- Harris, A. J., *How to Increase Reading Ability*. New York: McKay, 1965, 179-180.
- Heilman, A. W., *Principles and Practices of Teaching Reading*. Columbus: Merrill, 1967, 181.
- Johnson, M., and Kress, R., *Informal Reading Inventories*. Newark: International Reading Association, 1965, 8.
- Ladd, L., *A Comparison of Two Types of Training with Reference to Developing Skill in Diagnostic Oral Reading Testing*. Unpublished doctoral dissertation, Florida State University, 1961.
- Lockerbie, D., *The Relationships between Three Methods of Measuring Reading Ability*. Unpublished master's thesis, University of Saskatchewan, 1969.
- Mitchell, R., *A Comparison of Children's Responses to an Original and Experimental Form of Subtests G-S and ND of the Gates Basic Reading Tests*. Unpublished doctoral dissertation, University of Minnesota, 1967.
- Powell, W., Reappraising the Criteria for Interpreting Informal Reading Inventories. In D. DeBoer (Ed.), *Reading Evaluation*. Newark: International Reading Association, 1969, 100-109.
- Powell, W., The Validity of the Instructional Reading Level. In R. Leibert (Ed.), *Diagnostic Viewpoints in Reading*. Newark: International Reading Association, 1971, 121-131.
- Rankin, E. J., Jr., The Cloze Procedure—A Survey of Research. In E. L. Thorstone and L. Hafner (Eds.), *Fourteenth Yearbook of the National Reading Conference*, 1964, 14, 133-150.
- Rankin, E. J. Jr., and Culhane, J., Comparable Cloze and Multiple Choice Comprehension Test Scores. *Journal of Reading*, 1969, 13, 193-198.
- Spache, G. D., *Reading in the Elementary School*. Boston: Allyn and Bacon, 1969, 330-331.
- Weaver, W., and Bickley, A., Sources of Information for Responses to Reading Test Items. *Proceedings*, 75th Annual Convention, American Psychological Association, 1967, 2, 293-294.

JOHN P. MILLER

*The Ontario Institute for Studies
in Education*

The Effects of Human Relations Training On Teacher Interpersonal Skills

The study was undertaken to assess the effects of inservice human relations training on teacher interpersonal competence. The investigation was based on the assumption that the ability to increase significantly teacher interpersonal skills would also directly affect student development in a positive manner. The results of the study indicated the teachers gained significantly in interpersonal skills in a counselling relationship but that the acquired skills did not transfer to classroom interaction. The findings are discussed in a broader theoretical and empirical framework.

The children called upon us as persons, and we responded as persons. School was not a parenthesis inserted within life, but was actually an intensified part of life. It was just this, of course, that made the job so wearing for the teachers. But it was just this, too, that was the one source of every good thing that happened (Dennison, 1969, p. 33).

Some educators are arguing today that teachers must be able to respond openly with their students if any meaningful learning is to occur. The argument is not a new one. In the fourth century Augustine felt that learning occurs from the teacher-pupil relationship being a "mutual dwelling in each other", rather than teacher domination of student behavior.

In the nineteenth century Froebel articulated the need for the teacher's capacity to respond warmly and genuinely. Castle (1970) summarizes his position:

He reminds us that there lies deep in most human encounters in home and school a challenge that cannot be met with any slick response. And in this image of the teacher as the trustee of the child's capacity for self-direction, he reveals him as the supremely sensitive practitioner, not in the mechanical manipulation of teaching aids but in the intuitive control of the two-way traffic between maturity and childhood (p. 134).

Reminiscent of Augustine's "dwelling in each other" Martin Buber in this century developed the importance of the I-Thou relationship as a basic necessity to human growth. Finally, Abraham Maslow (1968) and the humanistic psychologists have argued that a "receptive" model of teaching is more conducive to growth than an "intrusive" model:

Above all, we would care for the child, that is enjoy him and his growth and self-actualization. So far this sounds much like the Rogerian therapist, his congruence, his openness and his caring. And indeed there is evidence by now that this "brings the child out," permits him to express and to act, and to experiment and even to make mistakes: to let himself be seen. Suitable feedback at this point, as in T-groups or basic encounter groups of non-directive counselling, then helps the child to discover what and who he is (p. 693).

In other words, teacher genuineness and empathy are conducive to the development of the student's own identity. There is now empirical evidence (Rogers, 1969; Truax and Carkhuff, 1965) to support the generalization that teacher empathy, respect and genuineness are positively correlated with student achievement and development. This investigation, then, was undertaken to assess the effects of a human relations program designed to increase the teacher's capacity to be empathic, respecting and genuine and was based on the assumption that the ability to increase teacher empathy would also directly affect student achievement.

Background to the Study

Two studies by Aspy (1965, 1967) are representative of the research that has been supportive of the stated assumption. In the earlier study Aspy studied the influence of teacher empathy, congruence, and regard on the reading achievement of 120 third-grade students. Aspy controlled for student ability by correlating the core conditions with high and low IQ students. The teachers tape-recorded their reading groups fifteen minutes each day during one week in March and again one week in May. The tape recordings were "blindly" analyzed by three experienced raters who employed the Truax and Carkhuff (1965) rating scales. The results indicated that high levels of teacher empathy, congruence and regard were related to higher student achievement on the Stanford Reading Achievement Test.

A later study by Aspy and Hadlock (1967) confirmed and extended Aspy's previous findings. Students taught by teachers high in empathy, warmth, and genuineness showed a reading achievement gain of 2.5 years during a five-month period while pupils taught by low condition teachers gained only 0.7 years. The truancy rate in classes with low-condition teachers was twice that occurring in high condition classrooms.

The findings of the two studies cited above as well as other studies (Christensen, 1960; Schmuck, 1963; Truax and Tatum, 1966, and Hefele, 1969) have indicated a correlative relationship only. Due to the inherent difficulties in controlling for all other factors affecting student development, a causal relationship cannot be asserted. However, since the evidence has been consistent with regard to the correlation described and due to the equivocal findings relating other teacher characteristics

(Grannis, 1970) to student achievement the ability to affect teacher empathy, respect and genuineness is valued as a constructive variable for teacher training.

The training selected for the study was the Carkhuff method (1969). It was selected over other types of human relations training for three reasons. In brief, these reasons included: first, the Carkhuff method focuses on the conditions conducive to student development; secondly, there is a reasonable amount of evidence to support the effectiveness of the Carkhuff training method with different samples, with approximately twenty studies (Carkhuff, 1969; pp. 301-310) indicating the overall effectiveness of the Carkhuff method; and, thirdly, the Carkhuff training method involves a structured format and is thus repeatable from one group to another. The Carkhuff training method is not unlike other methods in that small groups of eight to ten people discuss personal and/or professional problems under the guidance of a trainer. What differentiates it from T-grouping and encounter grouping is the employment of research scales which channel group interaction into one or more of the following dimensions:

1. empathic understanding
2. respect
3. genuineness
4. specificity of expression
5. self-disclosure
6. confrontation
7. immediacy.

The specific dimensions chosen for a particular training program depend on the group involved and the amount of time available for training. As to format, most Carkhuff training sessions will involve one group member acting as a helpee and one member as a helper. The helpee will either state something of real concern to himself or role play some type of situation. The helper in responding to him will attempt to communicate effectively using one or more of the scales to guide his responses. For example, in responding to the helpee's statement he might use the empathy scale to "respond with accuracy to all the person's deeper as well as surface feelings." After the helper has completed his response, the other members rate his response using the empathy scale as a guide for rating. Thus the scales form the basis for the two main elements in the Carkhuff program: communication training and discrimination training.

An excerpt from a Carkhuff session might elucidate the procedures. For example:

Trainee 1 acting as helpee: Uh, it's an empty life. It's, um, there is, uh, no depth to it at all. I mean you just talk about very, very superficial things, and the first few times, it's O.K. But then after that, there's nothing to talk about. So you drink and you pretend to be happy over silly jokes and silly things that people do when they all, uh, are trying to impress one another, and they're very materialistic, and uh, it's just not the route I want to go.

Trainee 2 acting as helper: So your feelings are so strong that you just can't fake it any more (Carkhuff, V. I, p. 219, 1969).

Trainer: O.K. How would you rate the helper's response on the empathy scale?

Trainee 3: I'd give him a three

Trainee 4: Three point five

Trainee 5: Three

Trainee 6: Four

Trainee 7: Three point five

Trainee 8: Three point five.

The trainer would then communicate his rating and provide the rationale for it. This pattern would then be repeated with other trainees acting as helper or helpee and using other rating scales to shape and discriminate responses. The trainer might also participate as a helper to offer a model for effective interpersonal functioning. (For further clarification, the response of Trainee 2 in the above excerpt is the communication phase of training and the trainees' ratings of his response is the discrimination phase.)

The trainer plays an important part in the training as he must be familiar with the scales and have developed a certain amount of expertise in implementing them for training purposes.

Method

Sixteen elementary school teachers volunteered to participate in the study. Since a training session can only accommodate 8-10 people the sample was necessarily small and thus the findings were limited. The sixteen teachers were randomly assigned to an experimental group or a control group. The experimental group received 18 hours of Carkhuff training held during an intensive weekend program. The trainer was experienced in the Carkhuff method and had been previously rated at a level on the Carkhuff scales necessary for trainer effectiveness. Due to the limited length of the program, the trainer concentrated on four dimensions: empathy, respect, genuineness and specificity of expression. Since there is more evidence indicating the applicability of these four dimensions to teaching effectiveness, they were selected for this particular training program.

Three instruments were administered. Two of the instruments assessed the effects of training in a one-to-one counseling relationship. The instruments included the communication index and the discrimination index which have reliably assessed communication and discrimination skills in a number of studies (Carkhuff 1969; pp. 94-132). The communication index contains sixteen written expressions to which the teachers wrote a helping response for each expression. As an example, one of the excerpts used in the index is:

Helpee: I don't know if I am right or wrong feeling the way I do. But I find myself withdrawing from people. I don't seem to socialize and play their stupid little games any more. I get upset and come home depressed and have headaches. It seems all so superficial. There was a time when I used to get along with everybody. Everybody said, "Isn't she wonderful. She gets along with everybody. Everybody likes her." I used to think that was something to be really proud of, but that was who I was at that time. I had no depth. I was what the crowd wanted me to be—the particular group I was with.

Each teacher response was then rated on a scale from one to five and the mean of these sixteen ratings formed the communication index. The discrimination index contains the same sixteen excerpts, each being followed by four responses: the teachers rated a total of 64 responses using the gross rating scale as a guide (Carkhuff, 1969, v. I, p. 115). The teachers' ratings were then compared with a key developed by experienced raters to determine the mean deviation from the key ratings (Carkhuff 1969, v. I, p. 127). Thus, an improvement in the discrimination index is indicated by a lower score or deviation from the key ratings. The third instrument was the Carkhuff rating scales; these were used to assess teacher empathy, respect, genuineness and specificity of expression in the classroom. The empathy scale runs from one to five. At the lowest level, "the first person appears completely unaware or ignorant of even the most conspicuous surface feelings of the other person. . . ." At the highest level the first person ". . . almost always responds with accurate empathic understanding to all of the other person's deeper feelings as well as surface feelings." Respect or positive regard involves a five point scale running from the lowest level where the ". . . first person is communicating clear negative regard for the second person . . ." to stage five where the first person "communicates a very deep respect for the second person . . ." The genuineness scale runs from level one where ". . . the first person's verbalizations are clearly unrelated to what he is feeling at the moment, or his only genuine responses are negative to what he is feeling as regards the second person . . ." to level five where the first person ". . . is freely and deeply himself in a nonexploitative relationship with the second person." Personally relevant concreteness or specificity of expression is defined by a five point scale where at the lowest level ". . . the first person leads or allows all discussion with the second person to deal only with vague and anonymous generalities . . ." and at the highest level the first person ". . . is always helpful in guiding discussion so that the second person may discuss fluently, directly and completely specific feelings and experiences . . ."

The teachers' classes were recorded for approximately fifteen minutes per day for an entire week so that there was a total time of an hour per week before and after the training session. Two three-minute samples were randomly excerpted from each one-hour tape. Thus, each teacher was rated on two segments, each three minutes in length before the training as well as two three-minute segments after training. This procedure was developed by Carkhuff (1969), and is based on evidence that reliability, range and discriminatory power of ratings are generally independent of segment length (Kiesler, Mathieu, and Klein, 1964). Each excerpt was then rated by two raters independently of each other and without knowledge of whether the ratings were pre- or post ratings.

In general, it was hypothesized that in comparison with the control group the experimental group would gain significantly in communication and discrimination skills in the counselling situation and manifest these skills in classroom interaction.

The Results

Generally, the results indicate that the training was effective in enhancing communication and discrimination skills in the counselling setting, since t-tests on mean gain scores were significant for both the indexes. These data can be found in Table I.

TABLE 1

	Pretraining Mean	Post-training Mean	t-test On Gain Score	P
Communication Index				
Experimental Group	2.20	2.78	2.26	.025
Control Group	2.09	2.15		
Discrimination Index*				
Experimental Group	1.13	.74	3.31	.01
Control Group	1.18	1.13		

* Low scores indicate *greater* discrimination ability.

The results, however, from the classroom data were negative. As the data in Table 2 indicate, the t-tests on the gain scores for the overall mean as well as on the individual dimensions (empathy, respect, genuineness, and specificity) were not significant. Indeed, all the means regressed from pretest to post-test. In the case of empathy the drop in score was significant for both groups.

TABLE 2

Means and Standard Deviations of the Pre- and Post-classroom Ratings for the Experimental and Control Groups

Dimensions	Experimental		Control	
	Pre	Post	Pre	Post
Overall Ratings				
Mean	2.29	2.10	2.05	1.87
S.D.	.51	.63	.38	.27
Empathy				
Mean	1.92	1.60	1.68	1.35
S.D.	.57	.50	.39	.17
Genuineness				
Mean	2.67	2.45	2.58	2.34
S.D.	.45	.55	.33	.31
Respect				
Mean	2.63	2.44	2.42	2.35
S.D.	.58	.31	.53	.35
Specificity				
Mean	1.94	1.88	1.52	1.45
S.D.	.77	.89	.53	.58

Discussion

The findings suggest the training was successful in enhancing communication and discrimination skills in a counselling setting but not in developing interpersonal skills for classroom interaction. This contrast concerning the effects of training is presented in Figure 1 where the communication index mean and the overall classroom rating mean are compared. One explanation of this phenomenon is that, although the training improved teacher interpersonal functioning, the school simply did not provide the environment for the change to take place.

The results are suggestive only, since the limitations of the study prevent any conclusive inferential statements from being drawn. A more reasonable approach is to refer to other literature which may help clarify alternative explanations for this finding.

Institutional Pressures and Interpersonal Relations

Philip Jackson (1968) from a number of observations carried out in elementary school classrooms has drawn some conclusions which may be helpful in understanding the findings of this study. Jackson's observations led him to the conclusion that "the dominant relationship in the classroom is quite impersonal when compared with that which goes on in the home" (p. 29). Because public school classes are often large, and because teachers are often concerned about classroom control, there is very little opportunity for teachers to establish meaningful relationships with their students. Adding to the impersonality, the teacher, in order to "teach effectively", may require that the students not communicate with one another except in a prescribed manner. As Jackson puts it, "In a sense, then, students must try to behave as if they were in solitude when in point of fact they are not" (p. 16). The classroom organization and institutional "climate" are therefore not conducive to teacher empathy.

Fuller's (1969) review of six empirical studies supports Jackson's conclusions that teacher emphasis on class control overrides other considerations:

As it is reported by these investigations, what we know is that beginning teachers are concerned about class control, about their own content adequacy, about the situations in which they teach and about evaluations by their supervisors, by their pupils and of their pupils by themselves.

The consistency of these findings is remarkable in the light of the different populations surveyed. The consistency lies not only in the similarity of concerns expressed but in the absence of concern about topics which are usually included in education courses: instructional design, methods of presenting subject matter, assessment of pupil learning dynamics of child behavior and so on (p. 210).

Similarly, McAulay (1966) reports from a survey of 648 elementary school teachers that eighty percent said they did not deal with controversial issues in their social studies programs. A common explanation for their reluctance was that the children get out of hand "when we discuss something unusual like that" (p. 29).

At a more general level, Argyris (1968) has pointed out how difficult it is to gain interpersonal competence when one's survival needs are high:

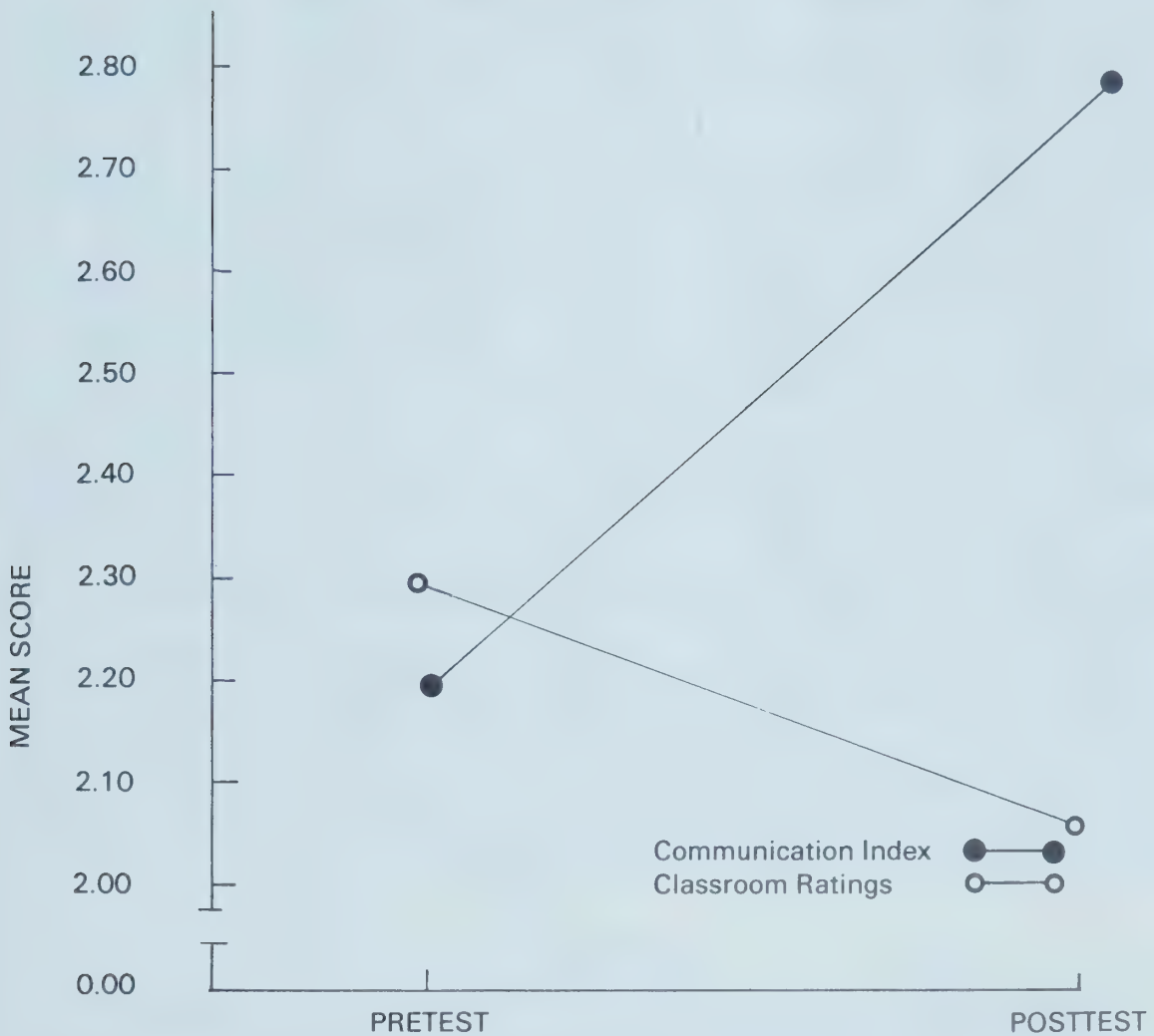
For example, the writer's interpersonal competence scores have been found to vary immensely depending upon the situation in which he was placed. Quantitatively his scores have ranged from 150 to 390, where the lowest score obtained is 10 and the highest 390. Interpersonal competence, therefore, is an interpersonal or situational ability and not simply an individual or personal ability (p. 149).

These data might support the conclusions drawn by Jackson and Fuller since the pressures of adequate lesson planning and the need to maintain discipline rate high as teacher "survival needs" and thus hinder teacher interpersonal competence.

Finally, Newmann and Oliver (1967) have suggested that the architectural structure and administrative design of the school may contribute to ineffective interpersonal relations. They point out that schools are often modeled after office buildings and factories where the students travel down corridors designed to handle traffic between different compartments. Administratively, the schools also resemble corporations with hierarchies of authority established with each individual (student, parent, teacher, principal, superintendent, and school committeeman) recognizing his duties within this framework. Activities are also organized into special categories: teaching (which is further compartmentalized into various subjects), administration, guidance, custodial services, etc. More particularly, education is packaged into equal and quantifiable units with classes of equal size, instructional periods of equal length, and standardized texts and lessons. In general, the institutional network that Newmann and Oliver describe may be another factor which could contribute to the inability of teachers to be empathic and genuine in the classroom.

The findings of this study are also similar to another study by Sutton (1968). In this study an experimental group experienced encounter grouping for a weekend and the effects were evaluated by the *Barret-Lennard Relationship Inventory*. The results were compared with two control groups. One control group consisted of teachers who volunteered for encounter grouping and a second group who did not volunteer. Neither control group experienced any form of training. Similar to the findings of this study, the three groups regressed with regard to empathy, congruence, and regard when assessed by the students. The regression that occurred over a five-week period was statistically significant for all groups. In other words, the students perceived the teachers as being less empathic after five weeks. Sutton attempted to explain this phenomenon by suggesting that the pretest measures were taken early in the year when the students' first impressions of the teachers were favourable. After five weeks, however, Sutton suggests these impressions must have faded. Perhaps a similar type of explanation could be made of the findings in this study. In short, the teachers' decline in interpersonal functioning could be interpreted as a sort of end-of-the-year phenomenon when teachers could be tired of teaching. In general, teachers at that time are probably more concerned with the coming summer than the final year-end demands.

FIGURE 1
Pretest and posttest mean scores on the communication index and the
classroom ratings for the experimental group.



The Training Methodology

It is also possible that the training was ineffectively designed to meet the needs of the classroom. Although the training was conducted with relevance to teaching skills, the fact remains that the training was done with adults performing the role of helpee. An alternative approach could incorporate microteaching situations with trainees rating each other on the Carkhuff scales after each microteaching incident. This would involve teachers working with students rather than with adults. In more general terms there is a need to develop training programs that focus on skills which are not removed from the social milieu where the trainees are working.

Finally, the training program may have simply been too short to

permit the skills to transfer to the classroom. Perhaps a training program over a longer period of time (six to twelve weeks) and constituting more time (thirty or more hours) would be more effective with feedback about classroom performance to the trainees.

*Implications of the Study for Teacher Education and
Suggestions for Further Research*

This study suggests that institutional pressures affect teacher interpersonal competence as well as the need for more data explicating the relationship between these two variables. For example, in view of Sutton's explanation it would be interesting to observe the effects of institutional pressures on teacher interpersonal functioning over an entire year. Perhaps certain "rhythms" could be detected so that inservice programs could be conducted at optimal times. In relation to the literature on institutional pressures (eg. Jackson; Newmann and Oliver; etc.) it would also be helpful to investigate the effects of school planning (eg. open vs. traditional) on teacher interpersonal functioning as well as the effects of class size, type of administrative network, and school architectural layout. It is interesting to note in this study that one of the two teachers in the experimental group whose level of interpersonal functioning *did* improve following on the training was teaching in an open plan school.

Finally, the data derived from the communication and discrimination index adds an additional piece to the large body of evidence supporting the effectiveness of the Carkhuff programs (Carkhuff, 1969) in improving counselling skills.

As noted earlier, the training needs modification if it is to affect teacher interpersonal skills in the classroom. However, even if a more effective program could be designed, it would be necessary to ascertain the effects of institutional pressures on interpersonal competence so that there would be greater conclusiveness concerning the probable outcome of training in different educational environments.

References

- Argyris, C. Conditions for competence acquisition. *Journal of Applied Behavioral Sciences*, 1968, 4, 147-177.
- Aspy, D. N. A study of three facilitative conditions and their relationship to the achievement of third grade students. Unpublished doctoral dissertation, University of Kentucky, 1965.
- Aspy, D. N. and Hadlock, W. The effects of high and low functioning teachers upon student performance. Unpublished manuscript, University of Florida, 1967.
- Barrett-Lennard, G. T. Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs*, 1962, 76 (Whole No. 562).
- Buber, M. I and thou. In Green, M. (ed.) *Existential encounters for teachers*. New York: Random House, 1967, 51-52.
- Carkhuff, R. R. *Helping and human relations*, 2 vols., Toronto: Holt, Rinehart and Winston, 1969.
- Castle, E. B. *The Teacher*. Toronto: Oxford University Press, 1970.
- Christensen, C. M. Relationships between pupil achievement, pupil affect-need, teacher warmth, and teacher permissiveness. *Journal of Educational Psychology*, 1960, 51, 169-174.

- Dennison, G. *The lives of children*, New York: Random House, 1969.
- Fuller, F. F. Concerns for teachers: A developmental conceptualization *American Educational Research Journal*, 1969, 6, 207-226.
- Grannis, J. C. The social studies teacher and research in teacher education. *Social Education*, 1970, 291-300.
- Hefele, T. J. The role of teacher interpersonal communication factors in the graduate education of prospective teachers of the deaf and the influence of such factors on the academic achievement of deaf students. Unpublished doctoral thesis, University of New York at Buffalo, 1969.
- Holder, B. T. Post-training retention effects as a function of training exposure. Unpublished doctoral dissertation, State University of New York at Buffalo, 1969.
- Jackson, P. *Life in Classrooms*, Toronto: Holt, Rinehart and Winston, 1968.
- Kiesler, D. J., Mathieu, P. and Klein, M. H. Sampling from the recorded therapy interview: A comparative study of different segments. *Journal of Consulting Psychology*, 1964, 28, 349-357.
- McAulay, J. D. Controversial issues in the social studies. *Education*, 1965, 86, 27-30.
- Maslow, A. Some educational implications of the humanistic psychologies. *Harvard Educational Review*, 1968, 38, 685-697.
- Newmann, F. and Oliver, D. Education and community. *Harvard Educational Review*, 1967, 37, 61-106.
- Rogers, C. R. *Freedom to Learn*. Columbus, Ohio: Charles E. Merrill, 1969.
- Schmuck, R. Some relationships of peer liking patterns in the classroom to pupil attitudes and achievement. *The School Review*, 1963, 71, 337-359.
- Truax, C. B. and Carkhuff, R. R. *Toward effective counselling and psychotherapy: Training and practice*—Chicago: Aldine Publishing Company, 1967.
- Truax, C. B. and Tatum, C. An extension from the effective psychotherapeutic model to constructive personality change in preschool children. *Childhood Education*, 1966, 42, 456-462.

HENRY L. JANZEN

DON BEEKEN

University of Alberta

and

JOHN HRITZUK

University of Calgary

Teacher Attitude as a Function of Locus of Control

Eighty teachers, attending an advanced educational psychology course were administered Rotter's I-E scale and Wehling and Charter's "Teacher Belief" questionnaire. Hypothesis predicted that those teachers scoring high on Rotter's "internal" scale would orient their belief system to "student autonomy", "integrative learning", "personal-adjustment ideology", "student challenge" and "consideration of student viewpoint". The "external" teachers were hypothesized to endorse "classroom order", "subject-matter emphasis" and "emotional disengagement". The statistical analysis did not confirm our hypotheses. In fact, the results showed the opposite effect. The "internal" teacher desired more control of his environment than did the "external" teacher. The study provides an explanation of these findings, forcing a reconsideration of some of the generalizations that have been made regarding locus of control. (Dr. H. L. Janzen is Assistant Professor in the Department of Educational Psychology, University of Alberta; Don Beeken is a graduate student in the Department of Elementary Education, University of Alberta, Dr. John Hritzuk is Associate Professor in the Department of Educational Psychology, University of Calgary).

The internal-external locus of control scale as developed by Rotter (1966) was an attempt to measure the extent to which a person believed that reinforcements were contingent upon one's own behavior ("internals") as opposed to being a mere result of fate, luck, or powerful outside forces ("externals").

A number of personal characteristics and expectations are often associated with the so-called internal or external individual. For example, Rotter (1966) suggests that an external individual perceives himself as relatively powerless to control his destiny. He feels more alienated from

the processes of his life and success and should therefore exhibit less achievement-oriented, striving behavior. The internal individual, on the other hand, believing that self-originated activity can be useful, is more willing to control his environment (or to exert the attempt). Rotter (1966) also suggests that an internally-oriented person should exhibit relatively more self-control. He may be more alert to his environment for data upon which to base his choices, and he should be more concerned with failure because he holds himself more responsible than would an external. The internal is also expected to see others more as he sees himself and to be considerate of their point of view.

The relationships mentioned above have persistently formed a basis for generating hypotheses, and explaining the results of research into locus of control and its relationship to other variables. It has not always been clear whether or not such statements are adequate and valid explanations.

Research relevant to the present investigation includes the work of Clouser and Hjelle (1970), who found a correlation between external locus of control and high scores on a dogmatism scale. External locus of control has been found to correlate positively with neuroticism (Platt, Pomeranz, and Eisenman, 1970), with manifest anxiety (Hountras and Scharf, 1970), and with blaming behavior (Phares, Wilson, Klyver, 1971). Lewis and Blanchard (1971) found that "internals" are resistant to subtle suggestions, and Bartel (1970) found that they attributed performance more to motivation than to ability. The results of a study by Weight (1970) suggest that "internals" are more confident in attributing good quality to an interpersonal relationship. Tseng (1970) found that they tended to make higher scores on scales which measured "compliance with rules", "ability to work with others", and "work tolerance". According to Hersch & Schiebe (1967), "internals" score higher on scales of dominance, tolerance, sociability, efficiency and well-being.

Problem

These findings, generalized to a teacher sample, suggested a number of hypotheses. It was expected that teachers classified as "internal" on Rotter's scale would also score high on measures of:

1. student autonomy
2. integrative learning
3. personal adjustment ideology
4. student challenge
5. consideration of student viewpoint.

It was predicted that teachers scoring as externals should score higher on attitude measures of:

1. classroom order
2. subject matter emphasis
3. emotional disengagement.

These attitude factors were derived from Wehling and Charters' (1969) questionnaire. The reliabilities of these factors were considered adequate (Wehling and Charters, 1969).

Method

Eighty teachers, attending an advanced psychology course at the University of Calgary completed Rotter's I-E scale and Wehling and Charters' Dimensions of Teacher Beliefs. The sample included teachers of elementary to senior high schools. The mean age was 35; average teacher experience was 5 years. The "belief" questionnaire provided measurements on the attitude scales referred to above. The eight dimensions of teacher belief were used as predictor variables of internal-external control.

Results

The teachers were classified as "internals" or "externals" on the basis of their scores. Subsequently a step-wise multiple regression analysis indicated that the following five factors, taken cumulatively, could be used to predict "externality" at a significance level of $p = 0.037$:

	Multiple R
1. student autonomy	.28
2. classroom order (1 + 2)	.30
3. personal-adjustment ideology (1 + 2 + 3)	.35
4. integrative learning (1 + 2 + 3 + 4)	.37
5. subject-matter emphasis (1 + 2 + 3 + 4 + 5)	.38

The results of this step-wise regression procedure, where variables below the 0.05 level of significance were *not* entered, failed to confirm the hypothesis that the measured attitudes of "student challenge" and "consideration of student viewpoint" were significant predictors of the "internal" teacher. In addition, the attitudes of "student autonomy", "integrative learning", and "personal adjustment ideology", believed to predict the "internal" teacher, in actual fact predicted the "external" teacher.

In the case of our second hypothesis, the attitude variables "classroom order" and "subject-matter emphasis" were found to predict "externality" as expected. The variable "emotional disengagement" was not a statistically significant predictor.

In order to clarify the extent of the relationships between the statistically significant attitude scales and Rotter's locus of control dimension, the following correlation matrix was obtained.

TABLE 1
CORRELATIONS OF ROTTER'S I-E SCALE WITH TEACHER ATTITUDES

	1	2	3	6	8	9
1 Subject Matter Emphasis	—					
2 Personal-Adjustment Ideology	0.02	—				
3 Student Autonomy	0.07	0.27*	—			
6 Classroom Order	0.26*	0.10	-0.24*	—		
8 Integrative Learning	0.12	0.55*	0.23*	0.31*	—	
9 I-E Scale	0.10	-0.01	0.30*	0.09	0.11	—

* significant at $p = 0.05$

The I-E scale is scored in such a way that the higher the score, the greater the tendency toward "externality". The correlation matrix indicates that the attitude variable "student autonomy" was the only one to correlate significantly with the I-E scale. The direction of the correlation shows that contrary to expectation, it is the "external" teacher who most endorses student autonomy.

Discussion

The significant correlation between subject-matter emphasis and classroom order (0.26) is in line with the expectation that a teacher who is concerned with curriculum would find an ordered class more suited to curriculum-oriented objectives. This finding is supported by the negative relationship between "classroom order" and the factor "student autonomy".

The results indicate significant statistical relationships between the dimension of "personal-adjustment ideology" and both "student autonomy" (0.27) and "integrated learning" (0.55). There is also a significant correlation between "student autonomy" and "integrated learning" (0.55), and between "student autonomy" and "integrated learning" (0.23). These three relationships form a rational and compatible group—possibly to be explained in terms of a concern for the individual child and his needs, especially in the area of developing responsibility. Such a concern would form part of an adjustment type of ideology in which the idea of encouraging student autonomy and free choice is seen as the basis of personal adjustment. Integrated learning not only correlates with the variables mentioned above ("personal-adjustment ideology", "student autonomy") but with classroom order (0.31).

Since the correlation between "student autonomy" and "externality" on the I-E scale is the main focus of this paper, an explanation of this relationship with reference to previous findings is warranted.

According to our findings, the "internal" teacher desires more control of his environment (in this case the classroom) than does an "external" teacher. This desire is manifest in his relative rejection of student autonomy. Oriented as he is to achievement, and being concerned with possible failure (Rotter, 1966, 1971), the "internal" teacher actively assumes responsibility for class control. Related to this would be a tendency to resist subtle coercion (Lewis & Blanchard, 1971), and any tendency on the part of the students to dominate, or weaken, the classroom environment. Because of the emphasis which the "internal" places on motivation, in contrast to ability, in explaining accomplishment (Bartel, 1970), we might expect him to consider student motivation as a very serious teacher responsibility. This attitude would not encourage any lessening of teacher responsibility by allowing student autonomy. The "internal" in general was found by Hersch & Schiebe (1967) to be dominant, assertive, independent, and efficient. These qualities would also be descriptive of a teacher who desires a controlled classroom. Such a teacher would probably be confident of his rapport with students (Weight, 1970) and unlikely to be neurotic (Platt, Pomeranz, Eisenman, 1970) or anxious (Hountras and Scharf, 1970). The "internal" is willing to comply with rules (Tseng, 1970) and shows a high score on the

Achievement via Conformance scale of the C.P.I. (Hersch & Schiebe, 1967). These qualities of compliance and conformity are not necessarily virtues: if manifested in a highly developed form they may indicate that the "internal" is less of an active innovator than we might hope, being less concerned with others than with his own situation. A tendency to be unable to see internal motivation in others could be one result of this lack of concern.

The "external" teacher on the other hand, believes that agents other than himself are responsible for life's events. We might assume that he feels more alienated from his destiny (Rotter, 1971). Less concerned with achievement, he is likely to accept failure calmly, secure in the belief that blame may be attributed outside himself (Rotter, 1966). Phares (1971) found that "externals" who had been "failed" on a standard test under normal conditions were more likely to exhibit blaming behavior (a defensive reaction) than in a condition which contained real distractions. Perhaps student autonomy functions analogously to the distractions in Phares' experiment: by allowing student autonomy he creates a less responsible and demanding situation for himself. The research of Lewis & Blanchard (1971) seems to support this hypothesis. The significance of this suggestion is that "externals" do not perceive subtly coercive situations as particularly threatening, indeed, they may welcome such situations. If flight from responsibility is characteristic of the "external," then one can easily explain his tendency to anxiety (Hountras, 1970) and neuroticism (Platt, et al., 1970). Indications that the "external" is less dominant, less assertive, more dependent, and more dogmatic (Hersch & Schiebe, 1967); Clouser & Hjelle, 1970) encourage the view that the "external" teacher is "pessimistic" about his personal power, fearful of responsibility, and consequently apathetic.

In summary: the "internal" teacher has been characterized as one who is concerned with, and believes in, his personal control. In contrast, the "external" has all but surrendered to "outside forces."

Both of these caricatures *seem unlikely* and it is our feeling that a closer consideration of the nature of locus of control could lead to a more realistic explanation.

Our findings indicate that the "external" teacher is aware of the importance of unpredictable or uncontrollable forces, and includes other individuals in this general view. Conceivably he is more aware of and alert to the environment than the "internal" is expected to be. No less than any other teacher, the "external" is aware that he holds a position of power in his classroom; however, it may be easier for him, not over-emphasizing self-assertion, to relinquish power and to consider the needs, goals, choices, basic individuality and freedom of others.

An internal locus of control manifests itself in a desire for control and in a strong sense of responsibility. These factors may make it more difficult for the internal to relinquish any control or possibly even to gain an objective viewpoint.

The present study forces us to reconsider some of the generalizations that have been made regarding locus of control. We are especially concerned with the question of whether the "internal" is actually more

willing and capable than the "external" of viewing others in the same light as he views himself. Is locus of control really representative of a philosophy of action or inaction? Does the individual who scores externally on the I-E scale really see himself as helpless and ineffectual, or is he expressing a belief in the nature of man in general—a belief which he may make use of in a very active and dynamic way? On the other hand, does an internal view also have such a generalizable philosophical basis or is it merely representative of a closed and personal belief in self-power? The results of our study seem to indicate that a passive, relatively unconcerned-with-others, "internal" is not an impossibility. Nor is an "external" who is active and innovative to be regarded as rare.

The notion that an "internal" would be more likely to view others as having qualities of internality, and therefore be more likely to respect their individuality and personal freedom, is seriously open to doubt. Our results indicate that the "internal" may be threatened by such a perception of others and react in a controlling and defensive manner. Perhaps further research would support the hypothesis that the "external" will be more accommodating in recognizing the individuality of others.

Clarification of the relationship between locus of control and activism is badly needed. Until we have some understanding of the reaction of internals and externals respectively to perceived freedom in others, "locus of control" will continue to be a poor predictor of interpersonal behavior, whether in teacher samples or in a more general class of subjects.

This study was supported by the University of Alberta's Research Grants and Contracts Office under Project 55-32481.

Bibliography

- Bartel, H. W. Attributions of Ability and Motivation in Person Perception and the Locus of Control of the Perceiver. *Dissertation Abstracts International*, 1970, Vol. 31, p. 462.
- Clouser, R. A. and Hjelle, L. A. Relationship Between Locus of Control and Dogmatism. *Psychological Reports*, 1970, Vol. 26, p. 1006.
- Hersch, P. D. and Schiebe, K. E. On the reliability and validity of internal-external control as a personality dimension. *Journal of Consulting Psychology*, 1967, Vol. 31, pp. 609-614.
- Hjelle, L. A. Internal-external Control as a Determinant of Academic Achievement. *Psychological Reports*, 1970, Vol. 26, p. 326.
- Hountras, P. T. and Scharf, M. C. Manifest Anxiety and Locus of Control of Low-achieving College Males. *Journal of Psychology*, 1970, Vol. 74, pp. 95-100.
- Joe, V. C. Review of the Internal-external Control Construct as a Personality Variable. *Psychological Reports*, 1971, pp. 619-640. Monograph Supplement 3-U28.
- Lewis, P. and Blanchard, E. B. Perception of Choice and Locus of Control. *Psychological Reports*, 1971, Vol. 28, pp. 67-70.
- MacDonald, A. P., Jr. Internal-external locus of control: A promising rehabilitation variable. *Journal of Counseling Psychology*, 1971, Vol. 18, pp. 111-116.
- Phares, E. J., Wilson, K. G., & Klyver, N. W. Internal-external control and the Attribution of Blame Under Neutral and Distractive Conditions. *Journal of Personality and Social Psychology*, 1971, Vol. 18, p. 285.

- Platt, J. J., Eisenman, R., & Darbes, A. Self-esteem and Internal-external Control: A Validation Study. *Psychological Reports*, 1970, Vol. 26, p. 162.
- Platt, J. J., Pomeranz, D., & Eisenman, R. Validation of the Eysenck Personality Inventory by the MMPI and the Internal-external Control Scale. *Journal of Clinical Psychology*, 1971, Vol. 27, pp. 104-105.
- Rotter, Julian B. External Control and Internal Control. *Psychology Today*, June, 1971, pp. 37-59.
- Rotter, Julian B. Generalized Expectancies for Internal Versus External Control of Reinforcement. *Psychological Monographs: General and Applied*, Vol. 80, No. 1, Whole No. 609, 1966 (#590895).
- Tseng, M. S. Locus of Control as a Determinant of Job Proficiency, Employability, and Training Satisfaction of Vocational Rehabilitation Clients. *Journal of Counseling Psychology*, 1970, Vol. 17, pp. 487-491.
- Wall, J. B. Relationship of Locus of Control to Self-Actualization. *Psychological Reports*, 1970, Vol. 27, p. 282.
- Weight, D. G. The Internal-external Locus of Control Dimension as a Variable in the Verbal Behavior of Subjects and Experimenters. *Dissertation Abstracts International*, 1970, Vol. 30, pp. 5246-5247.
- Wehling, L. J., & Charters, W. W., Jr. "Dimensions of Teacher Beliefs About the Teaching Process." *American Educational Research Journal*, 1969, 6, pp. 7-30.

JIM STAFFORD

The University of Alberta

Canadian School Life Tables

School life tables elucidate the relationship between age and school enrolment and allow us to present data in a standardized form for efficient comparison of "dropout" patterns between geographic areas. This paper provides us with the first set of Canadian school life tables and allows us to compare enrolment patterns in Canada, India and the United States. It reveals that Canadian children tend to begin their school life at a later age than U.S. or Indian children and that Canadian school participation rates peak a year earlier and thereafter decline more rapidly than U.S. rates.

I. Introduction

North American educators have spent a considerable amount of time and energy researching school dropouts, their causes, rates and consequences. They have found that dropouts correlate in varying degrees with intelligence (Tyler, 1956), academic performance (Havighurst, 1962), school attendance (Cerventes, 1965), grade failure (Hohol, 1955), discipline (Ede, 1967), socio-economic background (Hollingshead, 1949), geographic mobility (Zeran, 1962) and personality (Cohen, 1970). A basic consideration which has been overlooked in these studies is the incidence of dropouts by age that occurs at local or national levels. Such information would be valuable to educational planners whether they are concerned with the physical plant, the curriculum or with special programs. Planners at the governmental and post-secondary school levels would find data on patterns of secondary school dropouts by age and sex useful because such information combined with knowledge of the size of younger age cohorts would allow projections of future needs in post-secondary institutions as well as in the labour force.

A viable approach to the problem of detecting the incidence of dropouts involves tabulating the data in school life tables wherein a cohort of 100,000 births is followed through their school life span. Such a format would explicate the relationship between age and school enrolment. School life tables allow us to present data in a standardized form for efficient

comparison of dropout patterns between geographic areas or between periods of time in one area.

School life tables have been constructed for the United States by Stockwell and Nam (1963) and for India by Rao (1967) and by Premi and Chandra (1968), but none have previously been constructed for Canada. The procedure is similar to that used in constructing nuptiality (Mertens, 1965), labour force (Denton, 1969) and divorce tables (Krishnan, 1971). School life tables are a mixed-increment double-decrement type wherein the tables reflect gradual accession of the population into the school system and their separation due to mortality and dropouts. These two factors of attrition are treated in this table as being analytically distinguishable and mutually exclusive. Separation by mortality occurs when a person dies while attending school. Any other case of separation is classified as a dropout.

II. Methodology

The methodology in this paper replicates that of Stockwell and Nam in all but a few instances which will be pointed out later. The school participation of a large group of persons is traced from age four to age twenty-two, the youngest and oldest ages recorded for persons attending public schools in Canada during the period under investigation. As with conventional life tables, the school life table assumes an initial population of 100,000, although it documents only those who are attending. Members of the surviving population may enter school at various ages—usually 5, 6 or 7—and having done so, may leave at various ages due to mortality or dropout.

Two methods of determining life table rates include the “cohort” method and the “current” method. In the first technique a group of persons are followed through their life span and their ages at entrance to, and departure from, school are noted. The second method requires the tabulation of school participation rates and death rates at all ages in a given year. The school life table is then derived from these data. The latter method is most commonly used because of the availability of the data and the possibility of providing contemporary values. The results of the cohort method can only be compiled after the last individual in the cohort has left school.

Originally only 1961 enrolment data were used, but the derived participation rates were bimodal and could not be adequately smoothed using conventional smoothing techniques. Such procedures could be used to advantage once the average of the 1956 and 1961 rates had been taken. The danger of distorting the data is always present when smoothing formulas are used, particularly for curves which follow the pattern of school participation rates, experiencing rapid changes over a small age span. All standard smoothing techniques make use of values in the vicinity of the value to be smoothed under the assumption that these values change very little, and if they do change rapidly, they distort the value being smoothed. For this reason smoothing procedures should be avoided when possible and otherwise used with discretion to minimize their effect. This was accomplished by smoothing the school participation

rates as little as possible. A graduating formula was applied to ages 9 and 14 once, and 10 to 13 twice. The formula used was:

$$x_i' = 1/9(x_{i-2}) + 2/9(x_{i-1}) + 3/9(x_i) + 2/9(x_{i+1}) + 1/9(x_{i+2})$$

where x_i is the participation rate and x_i' the smoothed participation rate of age group i .

Such smoothing techniques were probably not used by Stockwell and Nam, who unfortunately do not describe the method they used.

III. Data and Findings

The school life tables in this paper are based on Canadian life tables, 1960-62, and on school participation rates for 1956 and 1961. These years were selected because they were the latest census years for which published data of school enrolments in Canada by age are available. The enrolment data refer to all elementary and secondary schools, which represent a wide cross-section of school systems in Canada because each province has the authority and responsibility for organizing its own education system while the federal government provides schools for Indian and Eskimo children and for children of National Defence personnel. For example, the number of grades a student was expected to pass through in 1956-61 varied from 11 in Newfoundland to 13 in Ontario, British Columbia and the Yukon. Furthermore, the availability of kindergartens differed from province to province, causing variation in the age at which students entered grade one. A second cause of regional variation was the provincial compulsory school attendance laws which required students to attend school during the ages 6-16 in Ontario, but only during the ages 7-14 in Saskatchewan. Such variations reduce the applicability of the ensuing school life tables to students in individual provinces and can only be validly applied to the Canadian school population as a whole.

This table can provide us with some insight into school life patterns of Canadian school children. School participation rates increase rapidly to a peak of 99 percent for persons 8-9 years of age, after which the rates decrease slowly until those ages when compulsory attendance laws become void during which the participation rates decrease rapidly.

The e_{sx} column, which depicts the expected number of school years remaining for persons who reach age x , reveals that between the ages of four and eight years, persons can expect to spend ten or eleven more years in school. The computed expectancies for these years change very little because only a small portion of the base population, 1_x , attends school before the age of seven. After age seven, the expected number of school years remaining decreases by approximately one year for each succeeding age cohort until age 18, at which age the school life expectancy is only 0.4 years. If the expected number of school years is converted to age at which separation from school occurs, it is found that persons between the ages of 7 and 17 can be expected to leave secondary school between the ages of 17 and 18.

It may be fruitful to compare this table with the 1957-59 United States school life table compiled by Stockwell and Nam and the 1960-62 table

TABLE 1
TABLES OF ELEMENTARY AND SECONDARY SCHOOL LIFE FOR THE TOTAL
POPULATION, 1956 AND 1961

Age	L_x	l_x	S_x	L_{sx}	T_{sx}	e_{sx}
4-5	96,878	96,915	.0025	242	1,060,296	10.940
5-6	96,814	96,845	.1189	11,511	1,060,054	10.946
6-7	96,756	96,784	.6477	62,669	1,048,543	10.834
7-8	96,706	96,730	.9701	93,814	985,874	10.192
8-9	96,660	96,682	.9877	95,471	892,060	9.227
9-10	96,619	96,639	.9805	94,735	796,589	8.243
10-11	96,581	96,639	.9751	94,176	701,854	7.266
11-12	96,542	96,600	.9733	93,964	607,678	6.293
12-13	96,502	96,561	.9700	93,607	513,714	5.322
13-14	96,460	96,523	.9547	92,090	420,107	4.354
14-15	96,410	96,482	.9359	90,230	328,017	3.401
15-16	96,352	96,437	.8661	83,450	237,787	2.467
16-17	96,285	96,383	.6905	66,485	154,337	1.602
17-18	96,208	96,321	.4997	48,075	87,852	0.913
18-19	96,124	96,248	.2725	26,194	39,777	0.414
19-20	96,034	96,168	.1005	9,651	13,583	0.141
20-21	95,937	96,988	.0264	2,533	3,932	0.041
21-22	95,834	95,887	.0146	1,399	1,399	0.015

*definitions of symbols are given in the Appendix.

for India compiled by Premi and Chandra. The methodology is similar in all three cases. One factor which detracts from the validity of the comparison is that the United States and Indian tables refer to persons enrolled in all educational institutions. Stockwell and Nam state that their enrolment figures refer to those in "regular" schools,

where a person may advance toward an elementary school certificate, high school diploma, or college or university degree. Persons attending other types of schools or courses are not regarded as enrolled unless credits are obtained which are transferable to a "regular" school. (Stockwell and Nam, 1963, p. 1114)

Premi and Chandra state that their figures cover "all types of education." (Premi and Chandra, 1968, p. 3.) The Canadian data in this paper refer to persons enrolled in elementary and secondary educational institutions. Nevertheless, comparisons of ages under 15 are justified since most of the United States and Indian school populations under 15 would be enrolled in elementary and secondary schools.

A second debilitating factor is the different smoothing methods involved in each of the three tables, a procedure which affects the derived values in the tables. Stockwell and Nam do not describe the methods they used. Premi and Chandra draw a free hand curve to smooth those ages beyond twenty. Assuming that all smoothing procedures attain the same result a comparison of the three school life tables for ages under 15 is viable.

Comparing the stationary school populations of the three countries,

we find the Canadian values are lower than United States values in each year while Indian values are lower than either except for the age 5-6 when the Canadian value is lower than that of India. The data reveal that Americans start their children in school at an earlier age than Indians or Canadians, that Indians rank second and that Canadians are most dilatory of the three countries in this regard.

TABLE 2
THE STATIONARY SCHOOL POPULATIONS OF CANADA, THE UNITED STATES AND INDIA

Age	Canada L_{sx}	United States L_{sx}	India L_{sx}
5-6	11,511	55,308	18,293
6-7	62,669	93,859	37,610
7-8	93,814	95,632	44,148
8-9	95,471	95,740	40,858
9-10	94,735	95,851	36,837
10-11	94,176	95,769	32,643
11-12	93,964	95,706	26,978
12-13	93,607	95,505	22,201
13-14	92,090	95,257	17,911
14-15	90,230	94,643	14,473

Indian participation rates fall rapidly behind those of the two North American countries for ages beyond 6 years, creating much lower values in the school life table. Canadian and United States values differ minimally from 8 to 10, but the American stationary school population is much larger than that of Canada before the age of 8. The stationary school population reaches its highest value in Canada in the interval 8-9, one year earlier than that of the United States. Attrition occurs more rapidly in Canada than in the United States suggesting that the school holding power is weaker in the former country. The difference in attrition rates is considerable and may have been created by differences in smoothing procedures or by the fact that the United States school population in these ages could have included some persons who were not attending elementary or secondary institutions.

Assuming that the differences in school life table values are a reflection of the school life patterns of the respective countries, we note in the preceding table a distinct difference in these patterns. Several possible factors that might explain the differences demand investigation. Such influences as compulsory school laws and propinquity of residence to school probably play a role in determining national school participation rates. Other factors might be the values attributed to education and the education cost/personal income ratio in the respective countries.

IV. Summary and Discussion

Educators have always expressed concern over the occurrence of school dropouts but they have never been provided with a standardized format for depicting the incidence of dropouts. In this paper the school life table method has been utilized because it meets the criterion of standardization, opening up possibilities for comparing the incidence of dropouts across time periods and across geographical areas.

Validating the data in this, as in most research, posed a problem. This problem is most apparent when one compiles life tables because all extraneous factors which create fluctuations in the data are removed. The remaining anomalies become starkly visible and are the effects of invalid data. Smoothing procedures were applied to school participation rates employed in the construction of the ensuing school life table and consequently they played a role in the portrayal of Canadian school life patterns. Consequently, a detailed description of these procedures is included in the appendix along with the general methodology employed in this paper, a methodology which replicates that used in the construction of school life tables drawn up by Stockwell and Nam for the United States and by Chandra and Premi for India. Endorsing equivalent methodologies facilitates international comparisons of school accessions and separations although the derived differences in national school life patterns are somewhat clouded by discrepancies in the definition of school participation. Nevertheless, striking differences between the three nations occurred at initial stages of school life. Canadian children began attending school later than both Indian and United States children. The incidence of school attendance by age in Canada quickly surpassed that of India but never achieved the same level as that of the United States while separations from Canadian schools occurred at a faster rate than from United States schools. Speculative explanations for such differences in school life patterns between Canada and the United States might involve the sparser population in Canada which necessitates a wider geographic distribution of schools. The different patterns may be a reflection of lower esteem for education, or of fewer economic incentives to prolong education in Canada.

The Canadian school population exists within twelve separate education systems, each with its own set of laws governing school attendance and each with its unique set of school life patterns. Construction of provincial school life tables would provide insights into regional differences within Canada and provincial tables for different periods of time would portray trends within these regions.

The work done in this paper was initiated by Dr. P. Krishnan of the Department of Sociology, University of Alberta. I have benefited from his many suggestions and criticisms.

References

- Cerventes, Lucius Ferdinand. *The Dropout: Causes and Cures*. Ann Arbor: University of Michigan Press, 1965.
- Cohen, Arthur M. and Florence B. Brawer. *Student Characteristics: Personality and Dropout Propensity*. Washington: ERIC Clearinghouse for Junior Colleges, 1970.
- Denton, Frank T. and Sylvia Ostry. *Working Life Tables for Canadian Males*. 1961 Census Monograph. Ottawa: The Queen's Printer, 1969.
- Dominion Bureau of Statistics. *Canadian Life Tables*, Cat. No. 84-510. Ottawa: The Queen's Printer, 1963a.
- Dominion Bureau of Statistics. *1956 Census of Canada*, Bul. 1-10. Ottawa: The Queen's Printer, 1957a.
- Dominion Bureau of Statistics. *1961 Census of Canada*, Bul. 1.2-3. Ottawa: The Queen's Printer, 1962.
- Dominion Bureau of Statistics. *Survey of Elementary and Secondary Education, 1955-1956*, Cat. 81-401. Ottawa: The Queen's Printer, 1957b.
- Dominion Bureau of Statistics. *Survey of Elementary and Secondary Education, 1960-1961*, Cat. 81-210. Ottawa: The Queen's Printer, 1963b.
- Ede, William John. "The Identification of Future School Dropouts by the Analysis of Elementary School Records." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1967.
- Havighurst, Robert J. *et al.* *Growing Up in River City*. New York: John Wiley and Sons, Inc., 1962.
- Hohol, Albert E. "Factors Associated with School Dropouts," *The Alberta Journal of Educational Research*, Vol. 1, No. 1, 1955, pp. 7-17.
- Hollingshead, August B. *Elmtown's Youth*. New York: John Wiley and Sons, Inc., 1949.
- Krishnan, P. "Divorce Tables for Females in the United States, 1960," *Journal of Marriage and the Family*, Vol. 33, No. 2, 1971, pp. 318-320.
- Mertens, Walter. "Methodological Aspects of the Construction of Nuptiality Tables," *Demography*, Vol. 2, 1965, pp. 319-348.
- Premi, Mahendra Kumar and Satish Chandra. "Tables of School Life for India," Paper read at the 1968 meetings of the Population Association of America, Boston, Mass., April 18-20, 1968.
- Rao, G. Nageswara. "A Note on the Life Table Model for Wastage Problems in Education," *Education and Psychology Review*, Vol. 7, 1967, pp. 189-196.
- Stockwell, Edward G. and Charles B. Nam. "Illustrative Tables of School Life," *American Statistical Association Journal*, Vol. 58, No. 304, 1963, pp. 1113-1124.
- Tyler, Leona E. *The Psychology of Human Differences*. New York: Appleton-Century-Crofts, Inc., 1956.

Appendix

The Appendix deals with the methodology used to compute the school life table. Table 1 depicts the table of school life for the total population in 1955-56 and 1960-61. It consists of conventional life table columns but with the values expressed in terms of school enrolment.

- L_x — the stationary population, representing the number of persons alive during each age interval x to $x+1$ assuming 100,000 births of each age cohort and assuming exposure to 1960-62 age-specific mortality rates in Canada for $4 \leq x \leq 22$. The stationary population is the number alive at the end of the interval (1_{x+1}) added to the average number of years lived by those who died during the

interval. Since it is assumed that mortality rates are constant during the interval x to $x+1$, the stationary population would be derived by:

$$L_x = \frac{1}{2}(l_x + l_{x-1})$$

l_x — refers to the number of males and females who would be alive at exact age x if 100,000 live births were subjected to 1960-62 age-specific mortality rates throughout their lifetime. The values were obtained by assuming a sex ratio of 105 males to 100 females at birth and by applying the formula:

$$l_x = (1.05 l_x^m + l_x^f) / 2.05$$

to 1960-62 life table data by sex, where l_x is the population of males and females who are alive at exact age x and l_x^m and l_x^f are the number of males and females respectively who reach age x as listed in the 1960-62 Canadian Life Tables.

s_x — enrolment rates obtained by averaging 1955-56 and 1960-61 rates and smoothing the rates which correspond to ages 9-14. s_x values for ages 4 and 5 are obtained by a different procedure because enrolment rates increase so rapidly during these ages. A free hand logistic curve was fitted to estimated rates at ages 4.0, 5.0, 6.5 and 7.5. From the curve, rates for tenths of a year for ages 4.0 to 6.0 were estimated and weighted according to the length of school participation time they entailed, and weighted participation rates were derived for each one-year age group.

L_{sx} — the number of persons who attend school during each age interval x to $x+1$ under the assumption that 100,000 annual live births are exposed throughout their lifetime to the age-specific mortality rates of 1960-62 and to the school participation rates derived from 1955-56 and 1960-61 data. Values were obtained by multiplying the L_x values by the corresponding school participation rates (s_x) in each age interval:

$$L_{sx} = L_x s_x$$

For the age interval 4-6, L_{sx} was obtained by multiplying the L_x value from the life table by the weighted participation rates.

T_{sx} — represents the total number of years in elementary and secondary school remaining to all persons enrolled in school. It is obtained by cumulatively summing the L_{sx} values from the oldest age to the youngest:

$$T_{sx} = \sum_x L_{sx}$$

e_{sx} — the mean number of school years remaining to persons who are alive at exact age x . It is calculated by dividing T_{sx} by the number of persons alive at exact age x :

$$e_{sx} = \frac{T_{sx}}{l_x}$$

TABLE A
POPULATION AND ENROLMENT DATA IN CANADA, 1956, 1961, AND
1956-61 COMBINED

CANADIAN POPULATION CANADIAN ENROLMENT PARTICIPATION RATES									
Age	1956	1961	Total 1956 and 1961	1956	1961	Total 1956 and 1961	1956	1961	Total 1956 and 1961
4	369,341	432,356	801,697	580	897	1,477	.0015	.0020	.0018
5	386,066	428,586	814,652	54,770	85,305	140,075	.1418	.2011	.1719
6	375,544	423,294	798,838	226,627	290,801	517,428	.6034	.6869	.6477
7	362,977	416,490	779,467	346,106	410,105	756,211	.9535	.9846	.9701
8	348,845	409,419	758,264	351,620	397,380	749,000	1.0079	.9705	.9877
9	333,621	401,733	735,354	356,084	384,128	740,212	1.0673	.9561	1.0066
10	317,524	394,116	711,640	303,606	376,955	680,561	.9561	.9564	.9563
11	300,776	387,090	687,866	284,582	366,710	651,292	.9461	.9473	.9468
12	285,113	376,177	661,290	282,165	362,900	645,065	.9896	.9647	.9754
13	271,506	359,444	630,950	274,144	361,693	635,837	1.0097	1.0062	1.0077
14	259,675	339,172	598,847	244,813	356,027	600,840	.9427	1.0496	1.0033
15	248,153	319,756	567,909	208,944	282,939	491,883	.8419	.8848	.8661
16	236,960	300,501	537,461	151,340	219,778	371,118	.6386	.7313	.6905
17	228,784	283,357	512,141	100,136	155,814	255,950	.4376	.5498	.4997
18	224,722	269,811	494,533	52,286	82,503	134,789	.2326	.3057	.2725
19	223,682	259,134	482,816	21,156	27,381	48,537	.0945	.1056	.1005
20	223,210	248,756	471,966	4,897	7,585	12,482	.0219	.0304	.0264
21	223,496	238,734	462,230	3,563	3,197	6,760	.0159	.0133	.0146

Source: DBS, 1956 Census of Canada, Bul. 1-10, Table 21.

DBS, 1961 Census of Canada, Bul. 1.2-3, Table 26.

DBS, Survey of Elementary and Secondary Education, 1955-56, Cat. 81-401, Tables 5, 54.

DBS, Survey of Elementary and Secondary Education, 1960-61, Cat. 81-210, Table 1-03.

TABLE B
SCHOOL PARTICIPATION RATES AND LIFE TABLE VALUES FOR
CANADA, 1955-1961

Age	l_x Both Sexes 1960-62	Participation Rates 1955-61
0	100,000	
1	97,269	
2	97,100	
3	96,997	
4	96,915	.0018
5	96,845	.1719
6	96,784	.6477
7	96,730	.9701
8	96,682	.9877
9	96,639	.9805*
10	96,600	.9751**
11	96,561	.9733**
12	96,523	.9700**
13	96,482	.9547**
14	96,437	.9359*
15	96,383	.8661
16	96,321	.6905
17	96,248	.4997
18	96,168	.2725
19	96,082	.1005
20	95,988	.0264
21	95,887	.0146
22	95,782	
23	95,673	
24	95,564	

* smoothed once
** smoothed twice
Source: DBS Canadian Life Tables, 1960-62, Cat. 84-510.

TABLE C
PARTICIPATION RATIOS FOR $4.0 \leq \text{AGE} \leq 5.9$ ACCORDING TO FREE
HAND LOGISTIC CURVE

Age	s_x	Length of Year Spent in School	Weighted Average
4.0	0.00	10	.0025
4.1	.00	9	
4.2	.00	8	
4.3	.00	7	
4.4	.00	6	
4.5	.00	5	.1189
4.6	.01	4	
4.7	.01	3	
4.8	.02	2	
4.9	.03	1	
5.0	.04	10	.1189
5.1	.06	9	
5.2	.09	8	
5.3	.11	7	
5.4	.14	6	
5.5	.17	5	.1189
5.6	.20	4	
5.7	.24	3	
5.8	.28	2	
5.9	.34	1	

JOHN McLEISH

University of Alberta

Teacher Education: Environments and Outcome

Students' opinions about their college environments are related to objectively determined features of their academic community; also to the actual measured changes which take place in these students on three year courses, shown by examination results, attitude and personality changes, values of a personal, social and religious kind. A complete sample ($n = 1,247$) of the intake and output of ten Colleges of Education in a single Institute of Education provided the raw data. The two main predictors of change in students were found to be the college emphasis on science and the democratic or "open" structure of the college.

This study falls into three parts: first, the attempt is made to measure students' attitudes to college environments using a subjective scale; second, an attempt is made to devise an objective index based on an inventory, using factual information such as size of college, qualifications of staff, etc.; third, an attempt is made to relate measured changes in students (in variables such as academic ability, educational attitudes, practical teaching) to the conditions of their environment as measured by the two scales, subjective and objective. The data analyzed relate to ten Colleges of Education associated in a single Institute of Education area in England.

I. The Subjective Evaluation Scale

In developing a scale concerned with the nature and quality of college environments, considerations of validity and reliability figure prominently. Preliminary inquiries about the validity of a rating scale method revealed that students appear to be incapable of using this method to characterize their environment (McLeish, 1969).

It seems to be a counsel of practical wisdom to begin where other research workers have left off. This meant collating existing measures and reports on college environments, amending scales already available for the purpose of developing one appropriate to the special circumstances of the colleges being investigated.

The *College Characteristics Index* (Pace and Stern, 1958) was used as a source of statements which could be suitably modified and reorganized to serve as a test-instrument. From various sources, including Pace and Stern, 100 favorable statements were written down descriptive of aspects of college environments. These were classified under ten general headings. Thus, for each of the ten dimensions, ten statements were available, every tenth statement referring to a particular dimension. Students are invited to agree or disagree with each of ten statements, using a standard record form. The total score on the environment index, obtained by counting the number of items with which students agree, provides a measure of the student's view of the total environment; the average score measures the attitude of the student group to a particular college environment. The scores on each of the ten dimensions provide a detailed picture of each college as perceived by the students.

After an item-analysis of the students' responses, and having discussed the scale in relation to their college with members of staff, the names of the dimensions were settled upon. These provide a description of what each group of statements purports to measure. The dimensions are named as follows:

1. Student Energy; 2. Concern for Individuality; 3. Social Commitment; 4. Staff Image; 5. Intellectual Climate; 6. Clarity and System; 7. Student Loyalty to College; 8. Humane Regulations; 9. Group Participation; 10. Lack of Tension. Full details, including the scale itself, are presented elsewhere (McLeish, 1970).

In constructing the scale only favorable statements were used, except for Dimension 10. This unidirectionality seems likely to produce a response "set" favorable to the colleges. But since the purpose was to compare the ten college environments and not to define them in an absolute sense, it was decided to proceed on the basis that *either* the response "set" would act in the same direction and be equally favorable to each college, *or* it would have the effect of spreading the average scores over a wider range without affecting the rank order. To test the possibility that students might simply accept the statements on the basis of response "set," a kind of "lie scale" was devised by reversing the direction of statements in the tenth dimension ("tension"). On analysis, it was found that students who accepted a great number of favorable statements also agreed with statements in the tenth dimension intended to be unfavorable. The reverse is also true. The hypothesis about "set" was thus verified.

II. The Comparison of College Environments

The scale was completed by two complete samples, consisting of 1,247 three-year students in ten colleges and 229 two-year "mature" students in the six of these colleges which accepted this kind of student, in the last term of their respective courses. The testing of the mature students was conducted as a "trial run," a year in advance of testing the main sample, to validate the instrument. It was of considerable interest to discover that about a quarter of the statements discriminated between the colleges. The rationale underlying the scoring scheme is that if two-thirds or more of the student body agree that a particular statement is true

or false about their college, it is probable that this consensus reflects the objective reality. If the agreement is less than two-thirds, the nearer it approximates to a 50-50 split, the greater is the probability that students are reacting on a random basis. In such a case, the responses are disregarded as being not relevant to that particular college. In other words, it is impossible to say whether the students do not know the facts, or whether their disagreement is a simple difference of opinion. In either event, the students' view, reflected in the average score given that statement, is considered to be non-evidential with reference to that particular college.

In view of the gratifying results of the trial run with the mature students, it was decided to use the scale in its original form. It was expected that the younger resident students, for whom the scale was originally designed, would experience little difficulty with any of the statements. This prediction was verified on testing exactly a year later.

The sample of 1,247 three-year students is considered to be the definitive sample of student opinion about the ten colleges. They had spent one, two or three years in residence; they had been in close touch with all the other year groups passing through at the same time as themselves; they were fully involved in college activities. These statements are not true of the mature students. On testing, the main sample had almost completed their courses, after three years of full-time study. It seemed they had nothing to lose by answering the questionnaire frankly. They were assured of the complete confidentiality of their replies, that the main object of the investigation was to delineate the realities of college life, and that the investigator's main interest was to discover how the quality of college life might be improved. The students understood from previous testing that completed scales were taken direct from the colleges and scored by research workers who were not concerned with them as individuals, and who had no particular association with any of the colleges.

The results of this testing operation indicated that the instrument really does discriminate—average scores for the different colleges range from 49 to 20, with a spread such that each college can be differentiated clearly from the others. A detailed analysis of the colleges and of the index is given in another place (McLeish, 1970). This paper will concentrate on providing further data to show the relation between these measures and "output" criteria in the form of changes produced in students during their college life.

If we intercorrelate scores on the ten dimensions ($n=1247$), all the correlations are positive, with the exception of the tenth dimension (lack of anxiety). Here the correlations are uniformly negative. (It should be remembered that this dimension is scored in a reverse direction to the other nine). This is true of every college separately, and of the colleges thrown together in one group. If we reverse the correlations on the tenth dimension, thus scoring all items in the same direction (favorable to the college), the intercorrelations can be accounted for in terms of two factors. The first of these arranges the ten dimensions in the order shown (Table 1).

This is clearly a general factor of *approval*. It accounts for 45 percent of the total variance. It tells us that the general excellence of a college

TABLE 1
FACTOR SATURATIONS OF THE COLLEGE ENVIRONMENT INDEX

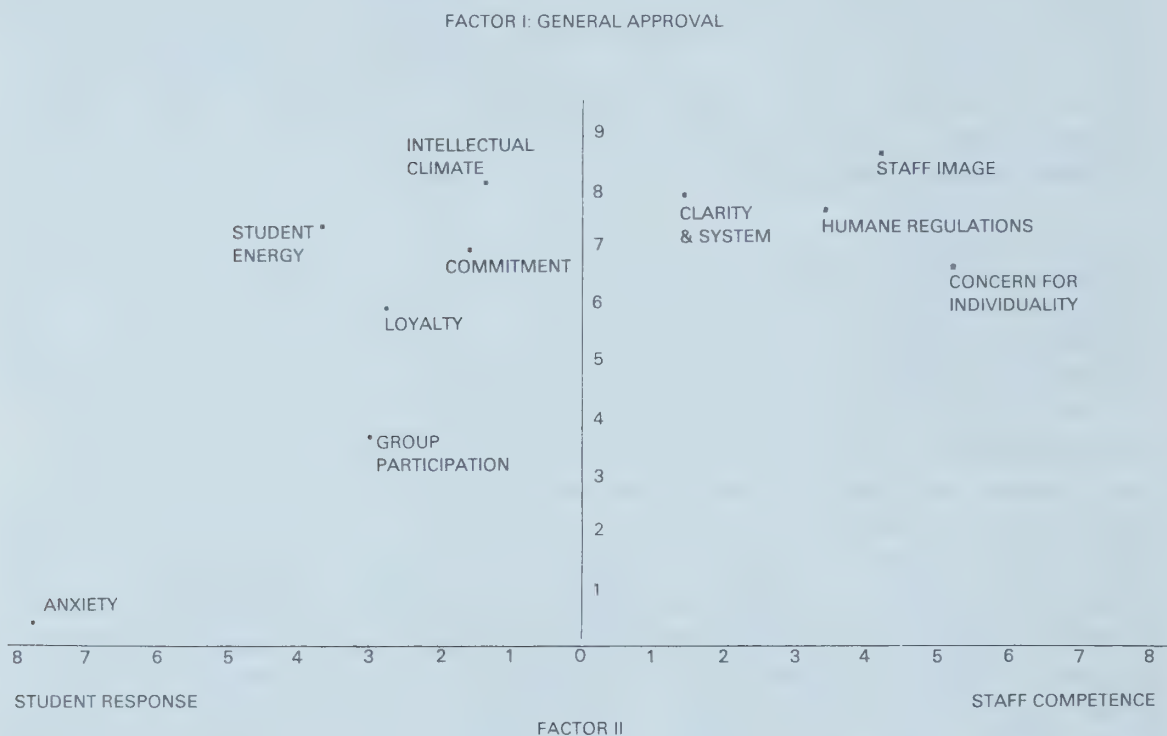
Dimension	Saturation Coefficients	
	First Factor	Second Factor
4. Staff Image	.77	.44
5. Concern for Individuality	.75	.54
6. Clarity and System	.74	.15
1. Student Energy	.74	-.30
8. Intellectual Climate	.72	-.12
3. Social Commitment	.71	-.17
2. Humane Regulations	.70	.36
7. Loyalty to College	.67	-.28
9. Group Participation	.46	-.30
10. Anxiety Level	.14	-.75

environment is seen from the student viewpoint as a function of a concerned and objective faculty, the concern for individuality shown by the college authorities, the clarity and systematic procedures with which courses are taught. Students also place considerable emphasis on the liveliness of the student body and the intellectual climate of the college. The concern of the staff with outside affairs (social commitment) and the loyalty of students to the institution come slightly lower down in the scale of priorities, but these features still function at a high level in the general pattern of approval. Group participation appears to be less important and the amount of anxiety generated by examinations as well as concern about career prospects appear to be quite unimportant, either positively or negatively.

On removing the influence of the first factor, the residual correlations yield a second, bipolar factor. This contrasts the ten dimensions in two groups. Intellectual climate, loyalty to college, social commitment, group participation, student energy, and anxiety are at one pole, whilst clarity-system, humane regulations, staff image and concern for individuality are at the other. This factor accounts for about 13 per cent of the total variance. It appears to contrast two kinds of environment: one end is characterized by the competence and humanity of the college staff, the other emphasizes effective student response to the environment. These relationships are set out diagrammatically in Figure 1.

Repeating the analysis, adding other variables consisting of the students' attitudes to the three teaching methods (lecture, tutorial and seminar) together with the thirty variables of the *Cambridge Survey of Educational Opinions* (McLeish, 1970), we discover several interesting clusters of variables. The ten dimensions of the College Environment Index are closely associated with each other in the first principal component. The general approval factor is now seen to be closely related to a stable conservatism in the student body: disapproval is associated with instability and radicalism. This dimension picks up also a tender-minded religious orienta-

FIGURE 1
The Inter-relations of the Ten Dimensions
of the College Environment Index



tion which seems to be associated with a favorable attitude to professional development, to the lecture method, and to teaching as a career. The first factor now accounts for only 12 per cent of the total variance most of this being accounted for by the *College Environment Index*. (Figure 2.)

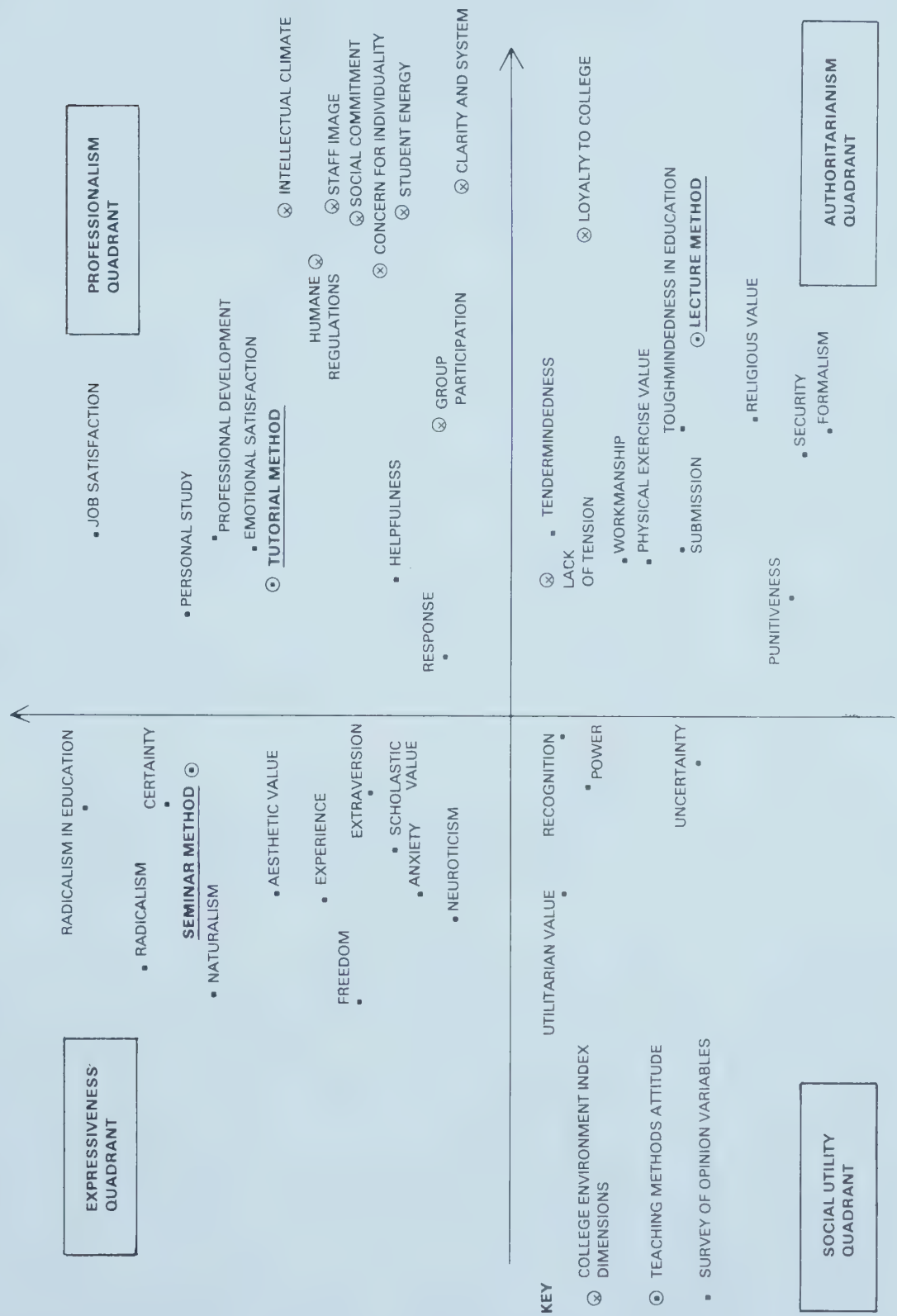
The second factor, which now accounts for about 9 per cent of the remaining variance, is clearly associated with a radicalism which goes with a favorable attitude to the seminar and tutorial. The two factors separate out: (i) a traditional, conservative, religious outlook favorable to the college, this being in contrast to a radical, naturalistic humanism which tends to be negative, even destructive in relation to teaching as a career and to the college provision; and (ii) a more discriminating radicalism which singles out particular aspects of the college environment as praiseworthy, which is committed to professional standards, and to teaching as a career.

In relation to these factors, it should be remembered that students as a group are not impressed by *any* of the colleges. In the best case, they accept only 49 per cent of the favorable statements as applicable to their environment; in the worst case only 20 per cent (McLeish, 1970).

Accepting that factors are best interpreted as principles of classification (Sir Cyril Burt has developed this idea at some length), these two orthogonal factors generate four quadrants, or pigeon-holes, within which the ten dimensions of the *College Environment Index* are disposed, along with the thirty-three other variables. In fact, the analysis shows that nine of the ten dimensions of the *Index* form a cluster within the same quadrant. Here too are located four variables related to satisfaction with teaching

FIGURE 2

Plot of College Environment Index with other Attitudinal and Personality Measures



as a career, two concerned with helpfulness and response to people, and another variable having to do with satisfaction with the tutorial method used in college teaching.

The *College Environment Index* variables, the values and attitudes of students and the measures of attitude to lecture, seminar and tutorial methods, judging by the way they group in the four quadrants, separate out four basic responses in teacher education—(a) a professional approach which catches up most of the favourable “sets” towards the college environment, (b) an expressiveness, or self-actualization approach which is associated with an unfavourable attitude towards teaching and teacher education, (c) an authoritarian, religious orientation favourable to the college environment and to the education provided, and (d) a set of attitudes and needs based on the criterion of social utility, these tending to be unfavourable towards the education and the environment provided.

III. The Objective Evaluation Scale

Although it is assumed that a consensus of two-thirds, or more, of the student body should point to some kind of reality, it is clear that a more objective measure is needed. A considerable amount of detailed factual information was available about each of the colleges. This was in the form of returns about the numbers and the training of staff, the numbers of students, and types of courses, etc. These returns were submitted periodically by each college to the Department of Education and Science. Pass lists for the final examinations (set by the Institute) were also available, as well as other materials. These factual data can be used to develop meaningful, quantitative indices which are obviously related to various aspects of the college environment. Twenty-three statistical indices were constructed to categorize the different colleges, using standard procedures (McLeish, 1970). The indices can be classified in four groups as follows:

1. Data relating to size of the College

Under this heading, five indices are available: number of students, of staff, rate of growth of the college, nature of the physical environment, the facility with which communication can be effected within the college system.

2. Data relating to character of the College

Here are several indices: the proportion of mature students; likewise the proportion of male students; the index of “masculinity,” being a measure based on the sex of the college principal and vice-principal, the proportion of male staff and of male students; “maturity,” based on the percentage of students over the age of twenty-five accepted for shortened two-year courses or for one year post-graduate courses; the index “community,” which seeks to answer the question: “How readily does the physical layout of the college lend itself to the development of an academic community?” and takes account of such things as proximity to a university or similar cultural centre, nearness to the metropolis, a reasonable layout of the college as an institution of higher learning; academic status which is determined from the pass lists processed by adding the passes, credits and distinctions which students obtained in the four parts of their final examina-

tion; library facilities, an index which takes into consideration the number, and the range, of books and periodicals in the college library; secular emphasis which was assessed in terms of the college emphasis on a narrow, fundamentalist religion, this being considered to be a handicap in an institution of higher learning. The staff-student ratio provides a measure of another important aspect of the college environment.

3. Data relating to Staff Qualifications

The professional qualifications of the staff were assessed in several ways. The main function of a college of education was taken to be the provision of a soundly based academic curriculum, with adequate and appropriate arrangements for the practical training of student teachers for the classroom. Staff scholarship is another variable in this group. It is distinguished from staff preparation, this consisting of the amount of practical training the lecturers themselves had experienced. Staff scholarship refers to academic training, it is calculated from the number on the faculty with graduate status and with advanced degrees. The staff training index was constructed on the basis of the proportion which had undergone teacher training, or taken advanced courses, or with a master's degree in education or some higher qualification in the area of teacher education. As a separate index the proportion completing an advanced full-time Institute course, or taking a master's degree in education in addition to their initial teacher training, was calculated. The ratio of untrained staff, as a percentage of total, provided still another index. In this case the scale was reversed, since a relatively large proportion of untrained staff was regarded as a disadvantage. Finally, under this main heading, a total impressionistic summary of "staff quality" was provided by the experimenter. In most of these cases a five-point scale was used.

4. Data relating to Courses

The nature of the courses provided constituted a fourth group of indices of significance in relation to an institution devoted to the initial training and education of teachers. The nature and diversity of courses constitute the basic data from which this index was constructed. In these colleges, courses fall in general into two groups: *academic* courses which are provided for personal development; and *professional* courses which provide for practical training. The latter are based on the notion of giving essential orientation to classroom activities whilst leaving students sufficient flexibility to adapt to the new and developing schools of tomorrow.

Indices were constructed as follows: the spread of courses dealing with teaching methods in the school subjects was assessed on a five-point scale; the main subject provided an index which summarized the extent to which colleges provided a spread of liberal studies in contrast to useful arts, crafts or practical training. The percentage of students in each college registered for the Bachelor of Education degree was used as an index of the diversification of courses. Similarly, the number of levels of training in terms of the age ranges of the children to be taught, was used as another index of diversification. Finally, the percentage of students studying for the advanced main paper in any subject provided another index of the quality of the college environment. The subjects

range through mathematics, the humanities, environmental studies, arts and crafts and advanced needlework.

The twenty-three basic indices were calculated on the same general plan. The colleges were assessed on a five-point scale according to the relationship each had to every other college on the index in question. It was assumed that on each index, *four* colleges could be given an average grade, *two* could be regarded as being above average, another *two* as below average; lastly, *one* college could be regarded as outstandingly above, and one other as outstandingly below average. No difficulty was experienced in transforming the numerical and other data into this five-point scale, the lines of division being quite sharp at the cutting-points.

One of the many advantages of this standardization technique is that it is possible very quickly to characterize the basic differences and similarities in the colleges. It is the object of the present study to discover, by appropriate analysis, these similarities and differences.

The impression derived from these indices is that in this group of colleges, one is outstandingly successful in developing an academic community dedicated to high scholastic and professional standards. At the other end of the scale we find a college where the history and situation, as well as the physical difficulties imposed by the site, militate against this kind of development. Between these two extremes we have a range of environments which differ in specific ways.

Covariance analysis can be used to discover the relationship between the colleges. There seems to be a hierarchical organization within the college group in terms of the quality of their environments. The best environment appears to be twice as good as the worst.

Comparing the two sets of indices (subjective and objective) it is possible to say whether, and to what extent the students' view of the environment is wholly subjective, without much relationship to those objective features which, it is generally assumed, predetermine the academic and humane quality of the environment (McLeish, 1970, pp. 103-109). Generally speaking, the students' impressions concur with the objective index.

IV. *The Relationship Between Environment and Outcome*

Data for each of the ten colleges on some sixty variables related to environment, plus fourteen variables related to educational product, were available. The first step consists of classification, with a view to reducing the number of variables. From scrutiny of the environmental variables it appeared that they could be placed in five categories, each of which singles out one aspect of the environment. These groups of variables were submitted to factor analysis separately to obtain *representative* variables, in the form of factors, for each category.

A. *Environmental Factors: Objective Indices*

Since factor scores were desired, the number of variables in each category was held to a maximum of nine, one less than the number of colleges. Excess variables were discarded according to two rules: (1) the variable correlated very highly with some other in the group; or (2) the

variable appeared to offer little promise of discriminating between the colleges. For example, whether or not provision was made by the college for the training of nursery teachers was discarded since only one college offered such training. This variable could therefore not be a basis for discriminating between the other nine colleges. Similarly, the number of students living at home was omitted since it correlated very highly (0.96) with the number of students on shortened courses. Likewise, the number of factors extracted from each group of variables was restricted to two in the interests of simplifying the analysis.

TABLE 2
Environmental Characteristics—Nature of Student Body

Variable	Factors		Communality
	I	II	
1. # of students (total)	<u>.717</u>	-.129	.531
2. % of male students (3 yr. program)	<u>.668</u>	.085	.453
3. % male students (2 yr. program)	<u>.894</u>	.240	.857
4. # of students (2 yr. program)	<u>.947</u>	.159	.921
5. % students in residence	<u>.04</u>	-.243	.876
6. % students in approved lodgings	-.394	-.370	.292
7. % students in infant training	-.203	<u>.671</u>	.492
8. % students in infant-secondary	.361	<u>.771</u>	.725
9. % students in secondary training	-.218	<u>-.942</u>	.934

On the basis of the relationships revealed by the Varimax rotations (see Tables 2 to 6), key environmental factors were identified. First of all, the colleges can be discriminated in terms of their student populations, the two factors in question being interpreted as "Coeducational-Student Maturity" and "Intended Teaching Level" (Factors I and II, Table 2).

The staff characteristics can be summarized most economically in terms of the factor identified as "Professional or classroom orientation versus

TABLE 3
Environmental Characteristics—Staff Variables

Variable	Factors		Communality
	I	II	
1. % graduate staff (full time)	-.313	<u>.596</u>	.454
2. % with teacher training (full time)	<u>.818</u>	.087	.677
3. % with advanced diploma (full time)	.303	.359	.221
4. # of full time staff	.385	<u>.531</u>	.431
5. Ratio part-time to full-time	<u>-.483</u>	.252	.297
6. % untrained grads (full time)	-.933	.002	.871
7. % untrained, no degree (full time)	<u>-.609</u>	-.025	.371
8. % male staff (full time)	-.158	<u>.941</u>	.911
9. Sex of principal (male)	.033	<u>.604</u>	.366

Personal (academic) development." The second factor is one of "Staff Masculinity" (Table 3).

The student evaluations of their college environments (Table 4) yields factors which can be labelled "General Quality of the Environment" and "Democratic or Open Structure," respectively.

TABLE 4
Environmental Characteristics—Student Evaluations

Variable	Factors		Communality
	I	II	
1. Facility of communication	-.159	.850	.747
2. Student energy	.883	.068	.785
3. Concern for individuality	.281	.828	.764
4. Social commitment	.798	-.092	.645
5. Staff image	.393	.865	.903
6. Intellectual Climate	.741	.458	.758
7. Clarity and system in courses	.759	.474	.802
8. Loyalty to college	.855	-.124	.746
9. Freedom from anxiety	-.148	.911	.853

The classification of students in terms of courses provides the two factors "Technical Sciences versus Humane Studies" and an "Academic versus Applied Emphasis" dimension (Table 5).

The last pair of factors (Table 6) catches up variables which are associated with the institutional features of the colleges. They can be appropriately labelled "Institutional Maturity" and "Institutional Utility" factors. The first of these factors contrasts the older, smaller foundations with more recent, rapidly expanding, more practically oriented colleges. The factor "Institutional Utility" implies that the college exists to meet a strong demand from the external environment. Such colleges are distinguished by having a large student body, they show a high rate of growth, they are relatively distant from a university, and are financed and supported by

TABLE 5
Environmental Characteristics—Students' Academic Programs

Variable	Factors		Communality
	I	II	
1. % studying humanities	-.888	.136	.808
2. % studying maths and sciences	.639	.654	.836
3. % studying environmental studies	.572	-.376	.469
4. % studying arts, crafts and music	.173	-.459	.240
5. % studying practical subjects	.938	-.219	.927
6. % taking advanced main subjects	-.238	.910	.885
7. % taking B.Ed. program	-.207	.860	.782
8. % studying divinity	-.543	.423	.473

a local educational authority. Although they tend to have a good library, their ratings on professional courses, and their evaluation by students is low.

TABLE 6
Environmental Characteristics—General Institution Features

Variable	Factors		Communality
	I	II	
1. L.E.A. or religious affiliation	<u>.626</u>	-.302	.482
2. square root of # of students	<u>.002</u>	<u>.810</u>	.655
3. character of library*	<u>.542</u>	<u>.806</u>	.944
4. rating of professional courses*	<u>.189</u>	-.582	.374
5. rate of growth (enrollment)	-.350	<u>.554</u>	.430
6. quality of physical environment*	<u>.897</u>	-.081	.812
7. rating of community life*	<u>.961</u>	.054	.926
8. total of student evaluation	<u>.616</u>	-.409	.547
9. distance from university	-.399	<u>.830</u>	.848

* These were rated on a five-point scale, with zero as the mid-point.

B. Product Factors: College Profiles

Having obtained scores for each of the colleges on each of these ten factors, attention was now focused on the fourteen product variables. These are: examination results, radicalism, punitiveness, formalism, naturalism in education, radicalism in education, religious value, utilitarian value, emotional, personal, professional and general satisfaction derived from teaching, tough-mindedness in education, anxiety, total change scores (pre-test to post-test) and the weighted score for change towards secular radicalism. (For a detailed description of variables and tests, see McLeish, 1970.) As a preliminary to factor analysis the scores on each of the product variables were standardized to a mean of fifty and a standard deviation of ten. Additionally, scores on those variables which had *negative* correlations with the total change score were reversed. The purpose of this analysis was to discover, if possible, colleges with similar product profiles. Analysis of the factor scores (Table 7) indicates that the colleges cluster in four groups. (College names are fictitious to conceal their identities: saturations which are underlined are taken to be significant.)

In each group, the classification is bipolar, certain colleges in each group being opposite in character to the others. In the first group, Mantague and Underwood are counterposed to Hathaway and Torridon. The profiles reveal that the colleges so grouped owe their similarities and differences to the dimensions Technical Sciences, Intended Teaching Level, Institutional Maturity, and Teacher Training. The second group, which includes Blandon and Pemberton, consists of polar opposites in respect of the Coeducational-Student Maturity, Intended Teaching Level and Technical Sciences dimensions. The third group, which includes Brookhurst

and Stanfield, are in contrast on *all* the environmental variables except for the Institutional Maturity dimension. Probably the most interesting finding is with respect to Group IV colleges in which Langley is opposed to Clarendon. One of the colleges has an environment which fluctuates at the extremes on the various environmental dimensions, being very high in some and very low in other measures, whereas the other college is consistently low on all ten environmental scales. That these two colleges should have similar educational outcomes is explained by an observation made by Pace and Stern to the effect that, as far as outcomes are concerned, it is about equally ineffective to have an environment which manifests incongruencies between the objectives and the demands made on the student, as to have one that is uniformly bad. Thus it is to be expected that an inconsistent environment (Pace and Stern's implicit press in conflict with explicit objectives) would have approximately the same effect on outcome variables as a relatively poor environment.

TABLE 7
Factor Analysis: Colleges Grouped by Factors

College	Group I	Group II	Group III	Group IV	Communi- ality
1. Blandon	.119	.935	.120	.096	.913
2. Brookhurst	-.041	-.156	.950	.017	.929
3. Clarendon	-.264	.034	.110	-.939	.964
4. Hathaway	-.769	.342	.061	-.452	.917
5. Langley	-.499	-.379	.386	.609	.913
6. Montague	.871	.022	-.144	.097	.789
7. Pemberton	.178	-.806	.039	.344	.801
8. Stanfield	.556	-.124	-.622	.027	.712
9. Torridon	-.683	.248	-.566	.246	.909
10. Underwood	.649	.122	-.052	.058	.443

To elucidate the relationship between environment and product, a final analysis was carried out on the transpose of the college-by-product matrix to identify characteristic dimensions of change. The matrix of correlations of the fourteen product variables was subjected to an hierarchical factor analytic procedure suggested by Henrickson and White (1966). In this way it was hoped to determine the relationship between each of the variables and the general change factor. Certain redundant variables were eliminated from this analysis at the outset. For example, the variables Utilitarian Value and Job Satisfaction were discarded because of their very low communality with the general change factor, only 0.025 and 0.101 respectively. Since Formalism correlated very highly (negatively) with Naturalism in Education, it too was dropped from further analyses. The scores for Total Change and Weighted Score for Change were also discarded, since these were linear combinations of other variables and hence redundant. In this way it was possible to limit the number of variables in the analyses to nine, as before. The varimax factor matrix from these nine

variables, together with loadings on the general change factor are shown in Table 8.

TABLE 8
Product Characteristics: Varimax Rotation

Variable	Factors I	II	III	IV	Commun- ality
1. Examination Result	-.076	.272	.831	.602	.771
2. Anxiety level (decrease)	.358	<u>.746</u>	.135	<u>.705</u>	.702
3. Radicalism (increase)	<u>.670</u>	-.131	<u>.654</u>	<u>.700</u>	.894
4. Punitiveness (decrease)	.007	<u>.837</u>	.055	.507	.704
5. Naturalism/Ed. (increase)	.664	<u>.720</u>	-.046	<u>.749</u>	.935
6. Radicalism/Ed. (increase)	-.138	<u>.554</u>	<u>.504</u>	<u>.530</u>	.580
7. Religious value (decrease)	<u>.920</u>	.070	-.038	<u>.547</u>	.853
8. Emotional Satisfaction (decrease)	.278	-.020	<u>.668</u>	<u>.546</u>	.524
9. Tough-mindedness/Ed. (decrease)	<u>.781</u>	.283	.264	<u>.766</u>	.760

Descriptive labels for these factors are "Radical Social Humanism," "Educational Permissiveness," "Academic Interest." The fourth or general factor is the general outcome, or total measured effect of the colleges on students' attitudes. Factor scores for each of the colleges on these four outcome measures are given in Table 9.

TABLE 9
College Environments: Product Outcome Measures

College	"Radical Humanism"	"Educational Permissive- ness"	"Academic Interest"	"General Outcome"	Total Sum (unweighted)
Blandon	54.1*M	55.5 M	64.5 H	63.8 H	237.9
Brookhurst	55.9 M	56.9 M	51.1 M	57.9 H	221.8
Clarendon	30.1 L	59.3 H	53.7 M	45.9 M	189.0
Hathaway	41.4 L	60.9 H	50.9 M	51.9 M	205.2
Langley	63.7 H	58.4 H	31.8 L	51.9 M	205.8
Montague	58.0 H	44.3 L	63.1 H	59.4 H	224.8
Pemberton	59.8 H	52.4 M	48.2 M	55.9 M	216.3
Stanfield	41.3 L	33.2 L	38.0 L	28.4 L	140.9
Torridon	43.3 M	46.9 M	41.4 L	39.2 L	170.8
Underwood	52.4 M	32.2 L	57.2 H	45.7 M	187.5

* H = relatively high; M = moderate; L = relatively low

Stepwise multiple regression (Draper & Smith, 1966), was used to determine which of the environmental dimensions could be related to each of the product dimensions. A probability level of 0.10 was found to be needed to identify a minimal number of predictors in this analysis. This suggests that the generalizability of the results is probably low: the

analyses should be viewed as exploratory (cf. McNemar, 1962, pp. 184-185). The correlation matrix shown in Table 10 provides an interesting series of hypotheses about the inter-relationships between outcomes and the nature of the college environment. It would appear that radical humanism is associated more or less closely with training for elementary schools in colleges where emphasis is placed more on personal development than on professional preparation, where there is an emphasis on humane studies rather than on scientific subjects and where there is a strong academic tradition and a large proportion of male staff (multiple correlation 0.73). Educational permissiveness seems to be a function of the system in general, all ten colleges producing this effect (this appearing as the most important single common objective in this group of colleges). It seems to be generated more efficiently where there is a co-educational student body, with a goodly proportion of mature students, in the relative absence of a professional orientation on the part of staff, which contains a relatively high proportion of males, and in the absence of an established academic tradition. Scholarship, or academic interest, appears as an outcome especially in relation to a democratic climate, coupled with an emphasis on academic versus applied courses (multiple correlation 0.62). This is true also of the general outcome measure which, of course, is heavily weighted towards humanism and a child-centered ideology.

TABLE 10
Environment x Product Correlations

Environment Dimensions	Product Dimensions			
	"Radical Humanism"	"Educational Permissiveness"	"Academic Interest"	"General Outcome"
1. Coeducational- Student Maturity	-.09	.34	.18	.25
2. Intended Teaching Level	-.55	.26	.24	-.03
3. Professional Orientation	-.48	-.30	.29	-.27
4. Staff Masculinity	.51	.10	.16	.45
5. Environment Quality	-.26	.34	.24	.19
6. Democratic Structure	.61	-.19	.59	.59
7. Technical Sciences	-.73	.30	.37	-.04
8. Academic Emphasis	.32	-.01	.56	.52
9. Institutional Maturity	.50	-.36	.39	.31
10. Institutional Utility	.11	-.04	.29	.21

Most of the conclusions to be drawn from the analyses have now been stated. Possibly the one which most needs to be reiterated is that

the results do not readily lend themselves to generalization beyond the data at hand. They should, however, facilitate the testing of suitable hypotheses using greater numbers of colleges. Alternatively, larger samples could be obtained by dividing particular colleges into smaller sub-cultures—especially in colleges which are large and diversified in the programs they offer.

The author wishes to thank Dan Precht for assistance with the computer analysis of the data and with the first draft of Part IV of this paper. Margaret Knight assisted in collecting, coding and analysing the data in the early stages of this inquiry. The University of Alberta provided funds and facilities for the computer analyses. The main study was carried out during the author's tenure as Research Fellow at the Cambridge Institute of Education.

References

- Astin, A. W. and Holland, J. L. The environmental assessment technique; a way to measure college environments. *J. Educ. Psychol.*, 1961, 52, 308-316.
- Burt, C. *The factors of the mind*. University of London Press, 1940.
- Draper, N. R. and Smith, H., *Applied regression analysis*, Wiley, New York, 1966.
- Henrickson, A. E. and White, P. O. A method for the rotation of higher order factors. *Brit. Journ. Stat. Psychol.* 1966, 19, 97-103.
- McLeish, J., The rating scale as a measure of college environments. *Journ. Res. in Educ. (Manchester)*, 1969, 1, 72-86.
- McLeish, J., *Students' attitudes and college environments*, University of Alberta and Cambridge Institute of Education, 1970.
- McLeish, J., College environments and student characteristics in *Teacher Education in Britain*. ed. D. Lomax, Wiley, 1973.
- McNemar, Q., *Psychological Statistics*, Wiley, New York, 1966 ed.
- Newcomb, T. M. and Feldman, K. A. *The impact of college on students*. San Francisco, Jossey-Bass, 1969, 2 vols.
- Nunnally, J. C., Thistlethwaite, D. L. and Sharon Wolfe. Factored scales for measuring characteristics of college environments. *Educ. Psychol. Meas.*, 1963, 23, 239-248.
- Pace, C. R. and Anne McFee. The college environment. *Rev. Educ. Res.*, 1960, 30, 311-320.
- Pace, C. R. and Stern, G. G. An approach to the measurement of psychological characteristics of college environments. *J. Educ. Psychol.*, 1958, 49, 269-279.
- Stern, G. G. Environments for learning. In Sanford, N. (ed.) *The American College*, New York, Wiley, 1962, 690-730.

REFERENCE BOOKS RECEIVED

INVENTORY OF CURRENT RESEARCH ON POST-SECONDARY EDUCATION, 1972

by Lon Hefferlin, Melvin Bloom, Jerry Gaff, Brenda Longacre
Center for Research and Development in Higher Education, University of California, Berkeley, 1972 xii + 291 (no price quoted)

Eleven hundred projects are listed by name of the research worker with a brief outline of objectives, institutions involved and materials already published. A comprehensive index provides guidance by cross-references to other projects on related topics. Coverage is confined to North America. The result is a useful and tastefully produced resource which will prove valuable to all research workers in this area.

PSYCHOLOGY IN THE U.S.S.R.: AN HISTORICAL PERSPECTIVE

by Josef Brožek and Dan Slobin
International Arts and Sciences Press 1972, x + 301 (no price quoted)

A collection of reprinted articles from Soviet sources with linking editorial matter provides a useful guide to sources of information in English and Russian. Topics discussed include learning, development, higher nervous activity, aviation and space psychology, military psychology, set, history from 1917, as well as accounts of where and by whom research is being conducted. The approach is encyclopedic with the virtue of comprehensiveness, and the associated defect of fragmentation. Emphasis on basic theory tends to be minimal, otherwise the editors have succeeded in their self-imposed task—to provide a working tool for those interested in obtaining access to Soviet work in a particular research area. The book should also interest those who need a picture of the general state of Soviet psychology and its special problems, theories and methodology.

THE TEACHER'S HANDBOOK

by Dwight W. Allen and Eli Siefman (eds.)
Scott, Foresman & Co., Glenview, Ill., xviii + 832, \$14.90

Eighty-seven leading American authorities contribute seventy-five articles on all aspects of the educational (teaching-learning) process. A comprehensive review of current American thought is presented for the beginning, and the experienced, teacher. The attempt is made to develop the concept of education based on some understanding of the perennial philosophical problems as well as the teacher's role in the developing and complex situations which confront us in the schools today. Annotated bibliographies accompany each article. The result is a pleasing and useful book which can be highly recommended for all those involved in teacher education.

CONTEMPORARY PROBLEMS IN HIGHER EDUCATION

by H. J. Butcher and Ernest Rudd (eds.)

McGraw-Hill Company, (U.K.), Maidenhead, England vi + 401, 7.20 (\$18.00).

Thirty authors review research and thinking about higher education in Britain over the past decade. The emphasis is on empirical research, most of which has appeared in British journals, or books, which are often relatively inaccessible. Inevitably, coverage is very uneven and topics discussed tend to lack coherence. The picture which emerges is of rapid growth in higher education and in research carried out piece-meal, motivated by the search for viable, if temporary solutions to administrative difficulties centering around inadequate financing, great growth in student numbers followed by a dropping-off, problems of student housing, unrest and disaffection amongst both students and faculty, problems of teaching method and student performance, as well as many others. The mixture is all too familiar and seems to be becoming more so in North America. The authors and editors are all actively associated with the (British) Society for Research into Higher Education.

DIRECTORY OF FACILITIES FOR THE LEARNING DISABLED AND HANDICAPPED

By Careth Ellingson and James Cass

My first reaction on examining this Directory was one of regret that so few Canadian facilities are listed. Whether or not this lack reflects a paucity of services for Canadian children with various handicapping disabilities should be the subject of investigation. One would suspect that many such services exist but are not included in this book.

The Directory covers ". . . diagnostic facilities and remedial therapeutic and developmental programs" for children with educational handicaps, speech and language disorders, physical disabilities, emotional disturbance and mental retardation. It is indexed according to states in the United States, provinces in Canada. The authors state that they obtained information from those facilities which were willing to provide data about their services. Each is then described according to the following criteria:

Diagnostic procedures; availability of specialists as consultants; remedial developmental and therapy programs; parent counselling. Funding sources are listed, as is the availability of the organization for training purposes.

In general, information included in the Directory about each institution appears to be relatively comprehensive. Only five Canadian provinces are listed as having facilities for the handicapped child: British Columbia, Quebec, Manitoba, Ontario, and Prince Edward Island. Alberta is not included. This omission leads one to question the usefulness of the Directory to Canadians interested in learning about services available in their own country. One wished, perhaps unfairly, that the authors had been somewhat more aggressive in their search for missing information. Notable U.S. facilities are also missing in some cases: however in view of the fact that submission of information was voluntary, this is perhaps understandable.

The Directory is certainly an interesting document, and its authors are to be commended for undertaking a massive task. It may well be a valuable first step in compiling a more complete index of services in the future.

Research in Education

Edited by F. MUSGROVE and H. J. BUTCHER

CONTENTS OF NUMBER 9, MAY 1973

Articles on

Successful Teachers

Children's Friendships

The Interview and Questionnaire

Classroom Observation

Teachers' Perceptions of Status Differences

Book Reviews

RESEARCH IN EDUCATION is published twice-yearly, in May and November, by Manchester University Press. Annual subscriptions are \$7.00 (post free); single copies are \$4.00.

MANCHESTER UNIVERSITY PRESS

316-324 Oxford Road
Manchester M13 9NR, England

